Diabetologia (2016) 59:2497–2502
DOI 10.1007/s00125-016-4044-y

MINI-REVIEW

# Key elements of metabolomics in the study of biomarkers of diabetes

**Jerzy Adamski** [1,2,3]

Received: 14 December 2015 / Accepted: 27 April 2016 / Published online: 6 October 2016
© Springer-Verlag Berlin Heidelberg 2016

**Abstract** Metabolomics is instrumental in the analysis of disease mechanisms and biomarkers of disease. The human metabolome is influenced by genetics and environmental interactions and reveals characteristic signatures of disease. Population studies with metabolomics require special study designs and care needs to be taken with pre-analytics. Gas chromatography coupled to mass spectrometry, liquid chromatography coupled to mass spectrometry or NMR are popular techniques used for metabolomic analyses in human cohorts. Metabolomics has been successfully used in the biomarker search for disease prediction and progression, for analyses of drug action and for the development of companion diagnostics. Several metabolites or metabolite classes identified by metabolomics have gained much attention in the field of diabetes research in the search for early disease detection, differentiation of progressor types and compliance with medication. This review summarises a presentation given at the 'New approaches beyond genetics' symposium at the 2015 annual meeting of the EASD. It is accompanied by another review from this symposium by Bernd Mayer (DOI: 10.1007/s00125-016-4032-2) and an overview by the Session Chair, Leif Groop (DOI: 10.1007/s00125-016-4014-4).

**Keywords** Biomarker · Diabetes · Mass spectrometry · Metabolism · Metabolomics · Nuclear magnetic resonance

**Abbreviations**

| | |
|---|---|
| FIA | Flow injection analysis |
| LC | Liquid chromatography |
| SOP | Standard operating procedure |

## Scope of metabolomics

The metabolites in a living system or a given sample are termed the metabolome [1]. Metabolites analysed by metabolomics are in the molecular mass range of 80–1200 Da. Metabolomics identifies a multitude (ideally all) of the metabolites in a given biological sample. In this way it provides a snapshot of the metabolites involved in distinct processes.

Whereas many functional features can be bioinformatically computed from the genome (e.g. RNA variants, splicing, protein sequences), the metabolome has to be analysed empirically and cannot be predicted from the genome. This is mainly due to the fact that metabolomics reflects input from the genome and the very dynamic environmental interaction with biochemical homeostasis. Furthermore, metabolomics is very closely linked to the functional phenotype, since the metabolites mirror dynamic processes that have been already performed or were happening at the moment of sample collection.

Metabolomics has been successfully used in the search for biomarkers for disease prediction and progression [2, 3], for analyses of drug action [4, 5] and for the development of companion diagnostics [6, 7]. Furthermore, metabolomics has been instrumental in discovering the impact of the

✉ Jerzy Adamski
adamski@helmholtz-muenchen.de

[1] Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Experimental Genetics, Genome Analysis Center, Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany

[2] German Center for Diabetes Research (DZD), Neuherberg, Germany

[3] Lehrstuhl für Experimentelle Genetik, Technische Universität München, Freising-Weihenstephan, Germany

genome on metabolic subtypes of human physiology [8, 9]. The metabolome in humans is influenced by sex [10], age [11], BMI [12], hormonal status [13], medication [14], nutrition [15], lifestyle (alcohol [16], smoking [17], coffee [18]) and diurnal rhythm [19], just to name the most penetrant confounders. However, the human metabolome is very stable over months [20] and even years [21] and deviations from conserved patterns may reflect a disease, environmental challenge or lifestyle change.

Several metabolites or metabolite classes have gained much attention in the field of diabetes research in the search for a method for early disease detection, differentiation of progressor types and compliance with medication. Among these metabolites, branched-chain amino acids [22], lysophosphatidylcholines, acylcarnitine and glycine [17], 2-amino adipic acid [23] and 1,5-anhydroglucitol [24] are being investigated for diagnostic clinical use. There are several excellent reviews dedicated to metabolic biomarkers in diabetes discovered by metabolomics [25–28].

## Generic processes in metabolomics

Metabolomics profits from the experience with other omics approaches, especially with respect to study design and biostatistics and bioinformatics. Table 1 depicts major elements of contemporary metabolomics. Projects in metabolomics have several requirements in terms of study design, since they generate a large amount of data [29]. Therefore, detailed documentation of phenotypes associated with samples has to be prepared and maintained. Human studies involve samples of urine, serum, plasma and saliva, which have very different metabolite spectra [25, 30–32]. The identification of controls appropriate to the aims of the project poses a special challenge. This is because of the fact that the human metabolome is influenced by many confounders with a known impact. In an ideal situation all samples in a large metabolomics experiment should be matched for confounders. As this is often not possible, a sample randomisation might improve the outcome of analytical procedures, and confounder effects must be corrected for during data analyses. Preparation and adherence to standard operating procedures (SOPs) is essential to maintain a cohort dataset as a sustainable resource. Pre-analytical issues could create a substantial component of variance prohibiting reliable biostatistical processing and interpretation of the data [33]. The common issues (e.g. multicentre study with different collection procedures, confused sample identity, insufficient sample volume) are known and should be avoided [34]. Handling of the liquid sample should include use of robots to improve both throughput and precision. Separation of metabolites prior to their identification has been proven to increase the resolution and sensitivity of analytics. Gas chromatography (GC) or liquid chromatography (LC) can be coupled to analytical instruments and improve the overall performance. Progress in analytical methods (especially in mass spectrometry) has made metabolomics possible and efficient in the last decade. Data evaluation is an integral part of metabolomics. Sophisticated biostatistics models are needed to interpret the collected data. As metabolomes of different species (e.g. human, plant or bacteria) are in part distinct but equally easily accessible in databases, care has to be taken to assign metabolites to pathways of health and disease. Quality control and quality management are essential in all steps of metabolomics, because without these procedures it is impossible to make decisions on outliers, metabolite identity and concentration bias/distribution and confidence intervals.

**Table 1** Essential elements of metabolomics experiments

| Element | Specific requirements |
|---|---|
| Study design | Large-scale data collection, control definition, randomisation, project documentation, quality management, SOPs, ethical issues, ethnicity |
| Pre-analytics | Sample collection, storage, transport and identity, adherence to SOPs, multicentre standardisation of SOPs, automatisation |
| Metabolite separation | Automatisation, stability, resolution, coverage, recovery, throughput, quality control |
| Analytics | Automatisation, stability, resolution, sensitivity, precision, range and limits of quantification, coverage, throughput, quality control |
| Data evaluation | Large-scale data computation, biostatistic models of evaluation, species-specific bioinformatics |

## Common metabolomics methods

There are a variety of analytical approaches that can successfully be used for metabolomics analysis [35, 36] (Table 2). The key element in the success of population-based metabolomics over the last decade is the availability of the quadrupole tandem mass spectrometer. Contemporary instruments are very robust, fast and sensitive, although they have a lower mass resolution. An MS/MS unit can be coupled to GC or LC to increase metabolite coverage. The GC-MS/MS requires laborious on-the-fly chemical derivatisation of metabolites prior to analysis. Because of the high temperatures in the GC unit, thermic labile metabolites cannot be identified. Highly polar metabolites are also identified by GC-MS/MS. Nevertheless, the GC-MS/MS is very popular in diagnostic laboratories for analyses of drugs and steroids and further

**Table 2** Analytics for metabolomics

| Approach | Metabolite coverage | Sensitivity | Throughput | Robustness | Present in clinical chemistry[a] | Apparatus cost | Sample cost |
|---|---|---|---|---|---|---|---|
| GC-MS/MS | Hundreds | High | High | Very good | Yes | Medium | Low |
| LC-MS/MS | Thousands | High | Very high | Very good | Yes | Medium | Low |
| LC-NMR | Few hundreds | Low | Medium | Extremely good | No | Very high | High |
| ELISA/RIA | Less than hundred | Very high | Low | Good | Yes | Low | High |

[a] These methods are already applied in clinical chemistry laboratories for the quantification of selected metabolites for diagnostic purposes

for fatty acids, sugars or tricarboxylic acid cycle metabolites in the discovery labs. LC-MS/MS does not require metabolite derivatisation (but, for example, amino acids benefit from this process) and allows the detection of a broad range of molecules (molecular mass <2000 Da) not covered by GC-MS/MS, such as amino acids, biogenic amines, organic acids and lipids. NMR or LC-NMR do not require any metabolite derivatisation and samples might be re-used for other analytics after the NMR analysis. Other advantages of NMR are its very high measurement stability and its resolution of lipids. However, NMR still reveals a major drawback in terms of sensitivity as only a few hundreds of metabolites can be quantified.

Two analytical approaches can be used for metabolomics: targeted and non-targeted. The features of these approaches are compared in Table 3. The experiments with GC-MS/MS are mostly targeted, whereas the LC-MS/MS could be performed either in targeted or non-targeted mode. These different approaches require distinct sample preparation and equipment tuning.

In the targeted mode only a select set of metabolites (often a complete metabolite family, e.g. eicosanoids) can be quantified during the MS/MS analysis. The simplest version of a targeted assay is the flow injection analysis (FIA), where the sample is directly injected into the mass spectrometer. FIA may work for many applications, including quantification of amino acids and lipids, but will not resolve certain isobaric compounds, such as leucine and isoleucine or lipids with the same total chain length. Therefore, in addition to FIA, further LC-MS/MS approaches are also popular. The analysis of the pre-selected metabolites is based on the characteristic fragmentation pattern that allows their unequivocal identification and quantification. For absolute quantification purposes, known concentrations of a set of internal standards with identical/similar chemical properties to the metabolites of interest (often isotopically labelled metabolites) are added to the sample and analysed together. Apparatus that is properly tuned and operating in the targeted mode can be very fast, robust and automated (Fig. 1).

In the non-targeted approach the analytical procedures are optimised to cover the entire metabolome present in the sample without focusing on a specific metabolite class. Logically, quantification is difficult as it is not possible to provide internal standards for all molecules. Instead, the differences in ion count for every metabolite analysed are used for semi-quantitative comparison. Non-targeted metabolomics requires instruments with high and very high mass accuracy to allow identification of the measured metabolites. This may include the use of quadrupole time-of-flight (Q-TOF), orbitrap, Fourier transform ion cyclotron resonance (FTICR), or quadrupole linear ion trap (Q-TRAP) instruments. The throughput and sensitivity in

**Table 3** Comparison of features of targeted and non-targeted analytics

| Feature | Targeted | Non-targeted |
|---|---|---|
| Metabolite coverage | Only selected | All possible |
| Number of metabolites | 10–200 | Thousands |
| Quantification | Absolute (e.g. μmol/l) | Comparative (e.g. fold change) |
| Processing speed | Very fast (e.g. 200 samples a day) | Slow (e.g. ten samples a day) |
| Comparability of results worldwide | Excellent | Limited |
| Stability | Excellent | Good |
| Running workload | Low | High |

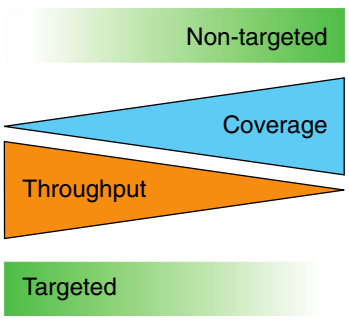


**Fig. 1** Comparison of throughput and coverage in targeted and non-targeted metabolomics. Specialised targeted approaches can be very fast but will not be able to provide comprehensive metabolome coverage. On the other hand, non-targeted approaches provide large coverage at slower throughput

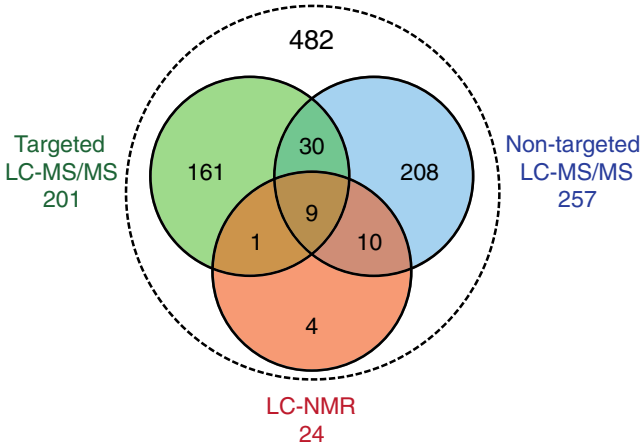


**Fig. 2** Comparison of the coverage of the metabolome with different methods. The number of all detected metabolites is given in the circle with the dotted line. Distinct metabolomic analytical methods reveal different unique but also common metabolites as indicated in the Venn diagram. Modified from [37]

the non-targeted mode is lower than that in the targeted mode. A common problem associated with metabolomics analysed by LC-MS/MS is that the mass spectra are hardware-dependent. Therefore, the same metabolite may have different features, such as retention time or fragmentation spectra. As a consequence, it is still a challenge to compare data from different sources containing unknown (not annotated) metabolites.

It has to be said that for the quantification of a small number of metabolites, any ELISA or RIA would be superior to mass spectrometry or NMR analyses in terms of sensitivity. However, the antibody-based quantification of metabolites has the drawbacks that it has low throughput and is limited to a couple of metabolites. Furthermore, antibody cross-reactivity limits the selectivity of assays.

If the same sample were to be subjected to different analyses available for metabolomics, some metabolites would be detected by only one approach, whereas others would be identified by more than one. In the example shown in Fig. 2 the serum sample underwent analyses by targeted LC-MS/MS, non-targeted LC-MS/MS and LC-NMR, and altogether 482 metabolites were detected [37]. In targeted and non-targeted LC-MS/MS 39 metabolites were found to overlap, and using all three approaches only nine metabolites (glucose, proline, alanine, valine, tyrosine, methionine, phenylalanine, histidine, lysine) overlapped. This example clearly shows that multiple approaches could be used to increase the coverage of the metabolome. Recently, human serum and urine metabolomes have been analysed with different mass spectrometry methods applied in parallel and revealed 4229 and 2206 metabolites, respectively [38, 39].

## Future developments

The area of metabolomics is developing very fast and several issues have been already identified as limitations. The study design may benefit from the rules regarding procedures already defined for clinical trials. The same applies to requirements for pre-analytical procedures, including collection, storage and transport. SOPs on pre-analytical procedures are present in public records but compliance is not high among different laboratories because there is no binding agreement on usage, or the elements of SOPs cannot be realised in the same way. Standardisation is a very large issue. Proposed approaches for standardisation, currently being investigated by many laboratories, include provision of reference substances and their mass spectra, as well as formats for data deposition in public repositories for metabolomics. In contrast to genomics or transcriptomics, metabolomics does not cover the whole metabolome. Therefore technological approaches, freely accessible mass spectrometrical analysis algorithms and data analysis resources (Table 4) are being developed worldwide to increase coverage of the metabolome.

**Table 4** Useful resources for metabolite analyses

| Resource | Focus | Link |
|---|---|---|
| HMDB (Human Metabolome Database) | Identity and biological description of metabolites, reference spectra of GC-MS, LC-MS and NMR, metabolic maps | www.hmdb.ca |
| LIPID MAPS | Nomenclature, structure and biological annotation of lipids | www.lipidmaps.org |
| KEGG (Kyoto Encyclopedia of Genes and Genomes) | Identity of genes and metabolites, description of pathways and reactions, chemical formulas | www.genome.jp/kegg |
| SMPDB (Small Molecule Pathway Database) | Annotated metabolic maps, data on disease involvement, drug characterisation | smpdb.ca |
| Standardisation (initiative by the Metabolomics Society) | Forum on different aspects of standardisation in metabolomics | www.metabolomics-msi.org/ |

**Acknowledgements** The author would like to express his gratitude for critical reading of the manuscript to J. Tokarz and M. Haid from the Helmholtz Zentrum München, Genome Analysis Center, Neuherberg, Germany.

**Funding** Work in the author's laboratory was supported by Innovative Medicines Initiative Joint Undertaking under grant agreement number 115439 (StemBANCC), and number 115317 (DIRECT), the German Federal Ministry of Education and Research (BMBF) to the German Center Diabetes Research (DZD e.V.) grant. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Duality of interest** The author declares that there is no duality of interest associated with this manuscript.

**Contribution statement** The author was the sole contributor to this paper.

## References

1. Tweeddale H, Notley-McRobb L, Ferenci T (1998) Effect of slow growth on metabolism of *Escherichia coli*, as revealed by global metabolite pool ("metabolome") analysis. J Bacteriol 180:5109–5116
2. Pena MJ, de Zeeuw D, Mischak H et al (2015) Prognostic clinical and molecular biomarkers of renal disease in type 2 diabetes. Nephrol Dial Transplant: Off Publ Eur Dial Transplant Assoc Eur Ren Assoc 30(Suppl 4):iv86-iv95
3. Wurtz P, Havulinna AS, Soininen P et al (2015) Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. Circulation 131:774–785
4. Janzer A, German NJ, Gonzalez-Herrera KN, Asara JM, Haigis MC, Struhl K (2014) Metformin and phenformin deplete tricarboxylic acid cycle and glycolytic intermediates during cell transformation and NTPs in cancer stem cells. Proc Natl Acad Sci U S A 111:10574–10579
5. Lord SR, Patel N, Liu D et al (2015) Neoadjuvant window studies of metformin and biomarker development for drugs targeting cancer metabolism. J Natl Cancer Inst Monogr 2015:81–86
6. Meadows NA, Morrison A, Brindley DA, Schuh A, Barker RW (2015) An evaluation of regulatory and commercial barriers to stratified medicine development and adoption. Pharmacogenomics J 15:6–12
7. Suhre K, Romisch-Margl W, de Angelis MH, Adamski J, Luippold G, Augustin R (2011) Identification of a potential biomarker for FABP4 inhibition: the power of lipidomics in preclinical drug testing. J Biomol Screen 16:467–475
8. Kastenmüller G, Raffler J, Gieger C, Suhre K (2015) Genetics of human metabolism: an update. Hum Mol Genet 24:R92–R101
9. Suhre K, Shin SY, Petersen AK et al (2011) Human metabolic individuality in biomedical and pharmaceutical research. Nature 477:54–60
10. Mittelstrass K, Ried JS, Yu Z et al (2011) Discovery of sexual dimorphisms in metabolic and genetic biomarkers. PLoS Genet 7:e1002215
11. Yu Z, Zhai G, Singmann P et al (2012) Human serum metabolic profiles are age dependent. Aging Cell 11:960–967
12. Floegel A, Wientzek A, Bachlechner U et al (2014) Linking diet, physical activity, cardiorespiratory fitness and obesity to serum metabolite networks: findings from a population-based study. Int J Obes (Lond) 38:1388–1396
13. Lankinen M, Schwab U, Seppanen-Laakso T et al (2011) Metabolomic analysis of plasma metabolites that may mediate effects of rye bread on satiety and weight maintenance in postmenopausal women. J Nutr 141:31–36
14. Altmaier E, Fobo G, Heier M et al (2014) Metabolomics approach reveals effects of antihypertensives and lipid-lowering drugs on the human metabolism. Eur J Epidemiol 29:325–336
15. Krug S, Kastenmüller G, Stuckler F et al (2012) The dynamic range of the human metabolome revealed by challenges. FASEB J 26:2607–2619
16. Jaremek M, Yu Z, Mangino M et al (2013) Alcohol-induced metabolomic differences in humans. Transl Psychiatry 3:e276
17. Wang-Sattler R, Yu Y, Mittelstrass K et al (2008) Metabolic profiling reveals distinct variations linked to nicotine consumption in humans—first results from the KORA study. PLoS ONE 3:e3863
18. Altmaier E, Kastenmüller G, Romisch-Margl W et al (2009) Variation in the human lipidome associated with coffee consumption as revealed by quantitative targeted metabolomics. Mol Nutr Food Res 53:1357–1365
19. Chua EC, Shui G, Lee IT et al (2013) Extensive diversity in circadian regulation of plasma lipids and evidence for different circadian metabolic phenotypes in humans. Proc Natl Acad Sci U S A 110:14468–14473
20. Floegel A, Drogan D, Wang-Sattler R et al (2011) Reliability of serum metabolite concentrations over a 4-month period using a targeted metabolomic approach. PLoS One 6:e21103
21. Yousri NA, Kastenmuller G, Gieger C et al (2014) Long term conservation of human metabolic phenotypes and link to heritability. Metabolomics 10:1005–1017
22. Wang TJ, Larson MG, Vasan RS et al (2011) Metabolite profiles and the risk of developing diabetes. Nat Med 17:448–453
23. Wang TJ, Ngo D, Psychogios N et al (2013) 2-Aminoadipic acid is a biomarker for diabetes risk. J Clin Invest 123:4309–4317
24. Yamanouchi T, Ogata N, Tagaya T et al (1996) Clinical usefulness of serum 1,5-anhydroglucitol in monitoring glycaemic control. Lancet 347:1514–1518
25. Yousri NA, Mook-Kanamori DO, Selim MM et al (2015) A systems view of type 2 diabetes-associated metabolic perturbations in saliva, blood and urine at different timescales of glycaemic control. Diabetologia 58:1855–1867
26. Park S, Sadanala KC, Kim EK (2015) A metabolomic approach to understanding the metabolic link between obesity and diabetes. Mol Cells 38:587–596
27. Suhre K (2014) Metabolic profiling in diabetes. J Endocrinol 221:R75–R85
28. Roberts LD, Koulman A, Griffin JL (2014) Towards metabolic biomarkers of insulin resistance and type 2 diabetes: progress from the metabolome. Lancet Diabetes Endocrinol 2:65–75
29. Tzoulaki I, Ebbels TM, Valdes A, Elliott P, Ioannidis JP (2014) Design and analysis of metabolomics studies in epidemiologic research: a primer on -omic technologies. Am J Epidemiol 180:129–139
30. Pena MJ, Lambers Heerspink HJ, Hellemons ME et al (2014) Urine and plasma metabolites predict the development of diabetic nephropathy in individuals with type 2 diabetes mellitus. Diabet Med: J Br Diabet Assoc 31:1138–1147
31. Yu Z, Kastenmuller G, He Y et al (2011) Differences between human plasma and serum metabolite profiles. PLoS One 6:e21230
32. Do KT, Kastenmuller G, Mook-Kanamori DO et al (2015) Network-based approach for analyzing intra- and interfluid metabolite associations in human blood, urine, and saliva. J Proteome Res 14:1183–1194

33. Plebani M (2006) Errors in clinical laboratories or errors in laboratory medicine? Clin Chem Lab Med 44:750–759
34. Fliniaux O, Gaillard G, Lion A, Cailleu D, Mesnard F, Betsou F (2011) Influence of common preanalytical variations on the metabolic profile of serum samples in biobanks. J Biomol NMR 51:457–465
35. Hyotylainen T (2012) Novel methodologies in metabolic profiling with a focus on molecular diagnostic applications. Expert Rev Mol Diagn 12:527–538
36. Artati A, Prehn C, Möller G, Adamski J (2012) Assay tools for metabolomics. In: Suhre K (ed) Genetics meets metabolomics: from experiment to systems biology. Springer, New York, pp 13–38
37. Suhre K, Meisinger C, Doring A et al (2010) Metabolic footprint of diabetes: a multiplatform metabolomics study in an epidemiological setting. PLoS One 5:e13953
38. Psychogios N, Hau DD, Peng J et al (2011) The human serum metabolome. PLoS One 6:e16957
39. Bouatra S, Aziat F, Mandal R et al (2013) The human urine metabolome. PLoS One 8:e73076