

Metabolomics for the rapid dereplication of bioactive compounds from natural sources

Nancy Dewi Yuliana · Muhammad Jahangir ·
Robert Verpoorte · Young Hae Choi



Received: 17 September 2012 / Accepted: 17 April 2013 / Published online: 25 April 2013
© Springer Science+Business Media Dordrecht 2013

Abstract The introduction of high throughput screening in the 1990s aimed to shorten the drug discovery route. To fully use its high potential, this technology requires a large number of compounds to screen. Thus, increasing the number of chemicals for initial screening is high on the agenda of pharmaceutical companies. High throughput synthesis and combinatorial chemistry were developed to address this demand. However, these technologies cannot fulfill the expectation to increase new lead compounds. One of the reasons is that the obtained compounds lack relevant chemical diversity. On the other hand, it is well known that an enormous molecular diversity and biological functionality are two important features which distinguish plant extracts as a drug source from synthetic chemicals, although a natural products-

based drug discovery project also poses some challenges, mostly connected with the presence of an active compound in a complex matrix with all kind of compounds. An elaborative purification to isolate and identify active compounds is thus needed. The possibility of antagonism or synergism between metabolites present in the extracts, and the fact that some common plant products have been found to be active in a number of test systems even make lead finding projects from natural sources become complicated. This review summarizes recent studies reporting metabolomics based techniques to uncover activity related compounds in complex plant matrices.

Keywords Chromatography · Drugs discovery · Medicinal plants · Metabolomics · NMR

N. D. Yuliana · M. Jahangir · R. Verpoorte · Y. H. Choi
Natural Products Laboratory, Institute of Biology, Leiden
University, Einsteinweg 55, 2300 RA Leiden,
The Netherlands

N. D. Yuliana (✉)
Southeast Asian Food and Agricultural Science
and Technology (SEAFAST) Center, Jl. Puspa 1,
Bogor Agricultural University, IPB Dramaga Campus,
Bogor 16680, Indonesia
e-mail: nancy_dewi@ipb.ac.id;
ndewiyuliana@yahoo.com

M. Jahangir
Department of Food Science and Technology, Hazara
University, Karakoram Highway, Dhodial, Pakistan

Introduction

Natural products lead finding

Natural products (NP) have historically been a rich source of lead molecules in drug discovery based on their capability to create unique and diverse chemical structures. The comparison of the major structural differences between combinatorial compounds, drugs, and NP molecules has been made. Natural products have a higher number of chiral centers, a higher number of oxygen atoms and a lower number of

nitrogen, sulfur and halogen containing groups than synthetic compounds. Natural products have a lower ratio of aromatic ring atoms to total heavy atoms, and a higher number of solvated hydrogen bond donors and acceptor functional groups (Feher and Schmidt 2003; Henkel et al. 1999; Newman and Cragg 2007; Newman et al. 2003). With these unique properties, in general, NP actually have more drug-like properties than synthetic compounds (Feher and Schmidt 2003; Koehn 2008).

Despite this clear advantage, supported by the fact that of all novel drugs in the past decades almost 50 % is a natural product or natural product derived compound (Newman and Cragg 2007, 2012; Newman et al. 2003), natural product lead discovery programs are not as common as one may expect. The major drawback is the supply. Being complex structures their synthesis is often too difficult to be commercially feasible, which means the production is dependent on the original natural source, which e.g. in the case of taxol was clearly a limiting factor for its development. But even already on the scale of screening for activity, plant material is many times not sufficient for a successful bioassay guided fractionation. That means sufficient yield for identification and structure elucidation as well as for testing biological activity. Another common problem is dereplication, i.e. to rapidly sort out if the activity is due to false positives or already known actives.

To address the problem of dereplication, one may consider the following approaches:

- Classical bioassay guided fraction
- Chromatographic screening for known actives or false positives
- Coupling separation with bioassay
- Systems approach in which metabolomics is used to rapidly link signals from the metabolomic analysis with activity.

Many studies reported possible approaches which could potentially help to overcome the above mentioned problem of known and false-positives in the NP drug discovery process, including the generation of a high quality NP compound library, the improvement of fractionation methods, and of dereplication and identification steps. Such studies involve the extensive use of various chromatographic techniques such as TLC, LC, and GC aiming at identification and quantification of (known) bioactives. In this approach

one tries to select samples for further analysis by excluding the known and false positive ones, but in fact this carries the risk that an unknown active is discarded when a well known active is present.

To couple separation with a bioassay is, thus, much more productive, it immediately links activity to signals, that means it measures real actives at the same time as the known and false positives. Several applications of coupling HPLC or TLC with bioassays have been reported, mostly involving various enzymatic bioassay systems (Ingkaninan et al. 2000; Rhee et al. 2001; Schenk et al. 2003; Schobel et al. 2001; van Elswijk et al. 2004). These techniques are useful to accelerate the finding of new leads from natural sources such as plants, microbes, or marine organisms. However, the compatibility of the assay systems to be incorporated into detection column is a prerequisite that may limit its applications.

The last mentioned option is a systems approach in which complex extracts can be tested in complex bioassays. By using different accessions, different extracts or fractions of an organism and measure the bioactivity of these samples one may be able to identify signals in the metabolomics data (chromatograms, mass- or NMR-spectra) that correlate to activity. These signals can then be linked to compounds. Advantage of this approach is that it can identify several active compounds in the same mixture, similar as in the chromatography coupled assays, but on top of that it may identify prodrugs and synergy in *in vivo* systems. Particularly in case of medicinal plants this is a major advantage.

Metabolomics

Different spectroscopic (NMR, MS, MS–MS) and chromatographic (HPLC, GC, GC–MS, LC–MS, TLC etc.) methods are widely used for metabolomic analysis of herbal medicines for quality control and lead finding (Shyur and Yang 2008). Nuclear magnetic resonance (NMR) is viewed as the most suited method for metabolomics based work because its excellent reproducibility, and thus suitability for public databases for data mining. Further advantages are simple and fast pre-analytical sample preparation, short measurement time, and the possibility to elucidate structures of known or unknown compounds in a complex mixture. However, its sensitivity (1 μM –1 mM in NMR tube) is not as high as MS and

chromatography based metabolomics platforms (Colquhoun 2007; Verpoorte et al. 2007), besides it is not available in every laboratory due to its high initial investment.

The main use of chromatographic techniques is to separate the complex extracts prior to detection. LC can be hyphenated with other techniques such as, diode array UV detection, (high resolution) mass spectroscopy (MS) and/or NMR. GC is commonly coupled with MS. Depending on the type of plant extract, GC–MS or LC–MS are the most common analytical tools, covering a broad range of metabolites (Badjakov et al. 2008). GC requires that the metabolites are volatile and thermostabile, which unfortunately excludes many important plant secondary metabolites such as phenolics, though derivatization at least enables the detection of many primary metabolites. Liquid chromatography is thus more suited to cover a large range of compounds present, though for LC the polarity window remains a limitation.

Chromatography-free techniques based on mass spectroscopy include different approaches of sampling such as direct infusion, laser desorption/ionization, imaging and some new ambient ionization methods may also be used for metabolomic study of mixtures (Han et al. 2009). The use of high resolution MS allows the identification of the molecular formulae of a compound, which is helpful in the identification of known compounds, though for every low molecular weight there are scores of compounds with the same molecular weight, not to speak about the number of novel un-known compounds that could have that molecular weight. Coupling MS with chromatographic tools (GC, LC, or CE) gives a further parameter that can be used to identify known compounds.

NMR-spectroscopy gives the most detailed information about the compounds present, and by applying various 2D NMR methods identification and structure elucidation in mixtures is many times feasible. Like in LC, in NMR the polarity of the solvent is a limitation. Solid phase NMR is an option, but peak broadening causes severe overlap of the signal. In any case the major advantage of NMR is that each proton gives the same signal intensity. Absolute quantification is thus simple, it only requires an internal standard, whereas in all other methods a calibration curve is required for each single compound, otherwise only relative values

for each single compound can be measured, but no quantitative comparison between peaks in the same chromatogram or mass spectrum is possible.

A very important aspect is the pre-analytical part of metabolomics. What you see is what you extract. What you extract depends on the method of harvesting, the quenching of all biochemical processes, the drying, the grinding and finally the method of extraction (Kim et al. 2010; Kim and Verpoorte 2010; Maltese et al. 2009; Verpoorte et al. 2008). Considering the extraction one should keep in mind that in any organism there is a wide variety of compounds present that differs greatly in physical properties, and that in any solvent will have a maximal solubility. There is no single solvent that dissolves everything. That means that there always be certain compounds that are at saturation level in the extraction solvent. Yuliana et al. (2011b) developed a new extraction method namely comprehensive extraction (Yuliana et al. 2011b). It is based on the continuous extraction of plant material with a mixture of solvents of increasing polarity (*n*-hexane, acetone, water) and the collection of fractions at predetermined intervals. In comprehensive extraction, the solvent is delivered directly into the column filled with powdered plant material mixed with Kieselguhr to remove void volume. The method allows the extraction of a wide range of metabolites with good resolution. Non-polar to polar compounds are eluted subsequently according to solvent gradient as can be seen in the PCA score plot (Fig. 1), resulting in more clear spectra of chemical profile of the plant material which are easier to interpret only by ^1H NMR spectra, or the NMR data can be subjected further to metabolite-bioactivity correlation study by using multivariate data analysis such as OPLS (Orthogonal Projection to Latent Structure). Thus, compounds correlated to the respected bioactivity can be directly identified from the extracts.

Standard protocols are thus an absolute requirement for metabolomics to be able to obtain reproducible results and to be able to separate the biological variation from the variation due to the analytical method. The different metabolomic approaches are all needed to have an as complete picture as possible of the metabolome of an organism, and direct comparison between these methods can only be made if exactly the same pre-analytical treatments are used, which usually is not the case as the sample that will be injected in LC, GC or measured in the NMR will be a

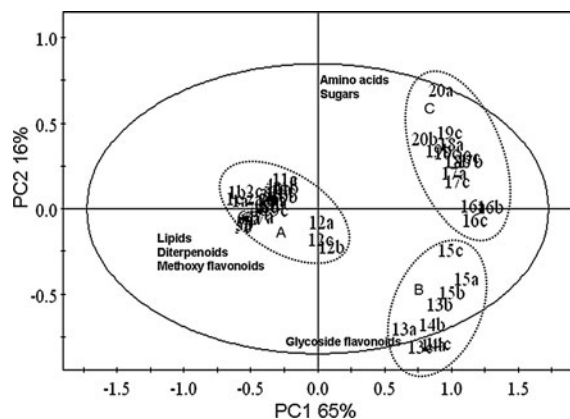


Fig. 1 PCA score plot of *Orthosiphon stamineus* fractions obtained from comprehensive extraction using combination of nonpolar to polar solvent in gradient (*n*-hexane–acetone–water). Numbers represent samples number while alphabets represent replications. **a** Group of fractions where nonpolar metabolites such as lipids, diterpenoids and methoxy flavonoids are predominant **b** Group of fractions where semi-polar metabolites such as glycoside flavonoids are predominant **c** Group of fractions where polar metabolites such as sugars and amino acids are predominant

different solvent, and thus a different polarity window. But by these approaches a variety of compounds (e.g. terpenoids flavonoids, alkaloids, glucosinolates, fatty acids, sugars, etc.) can be observed in metabolomic studies (Steinmann and Ganzera 2011).

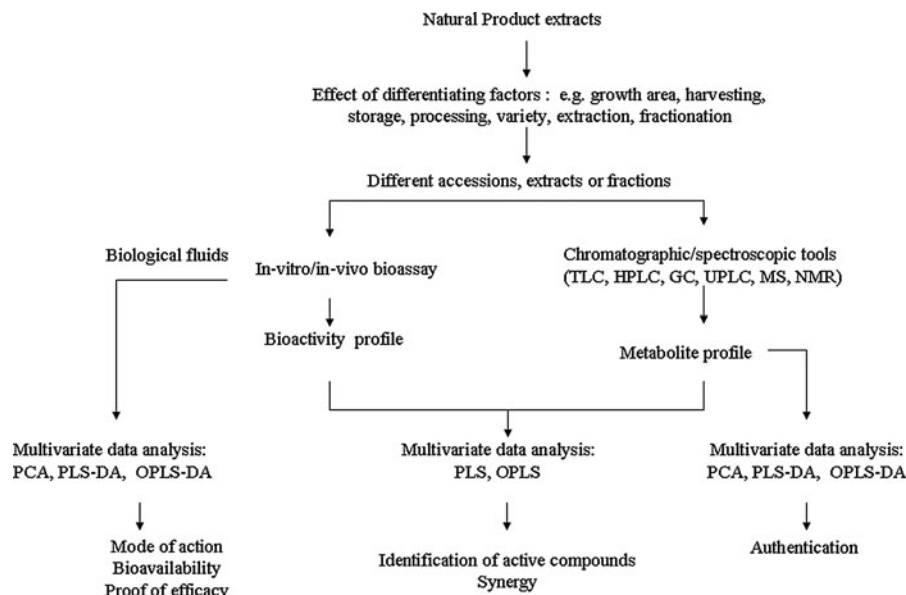
This review discusses the application of the aforementioned chromatography and NMR techniques in

metabolomics-based studies for quality control of herbal medicine and for identification of the biological active compounds in natural product extracts. A summary of methods described in this review can also be seen in Fig. 2.

Quality control of botanicals

Different metabolite profiles of natural product extracts may occur due to variations between species or variety, adulteration, environmental changes during growth and harvesting, post harvesting treatment, extraction and method of preparation. These factors may significantly alter the bioactivity profile of the extracts (Chang et al. 2011; van der Kooy et al. 2009; Wang et al. 2005). Therefore, for a safe and effective use of botanicals, a thorough quality control is required. In case of a well known active compound, a targeted approach will be the first choice, but for most medicinal plants more than one active compound is present, and even worse in most cases no active compound is yet known. The use of an unbiased method of analysis, i.e. metabolomics is thus a way to define quality. Van der Kooy et al. (2009) reviewed the application of metabolomics for this goal. Here we will give some examples to illustrate this sort of applications.

Fig. 2 Flow chart of metabolomics based lead finding from natural products



The use of HPLC fingerprints followed by application of PCA and PLS-DA was applied to examine the authenticity of the traditional Chinese medicines *Pericarpium Citri Reticulatae* and *Pericarpium Citri Reticulatae Viride* which contain *Citrus reticulata* ‘Chachi’, *Citrus reticulata* ‘Dahongpao’ and *Citrus erythrostris* Tanaka as well that of commercial preparations containing a mixture of many tangerine peels which had been made with new tangerine mutations of uncertain quality (Yi et al. 2007). Ma et al. (2008) reported an example of application of GC and GC–MS combined with multivariate data analysis to correlate artemisinin concentration with the growth stage of *Artemisia annua* L. and to distinguish between transgenic and wild type.

Ultra performance liquid chromatography (UPLC) emerges as a fast chromatographic method, with excellent peak shapes, and high reproducibility for the analysis of complex biological samples, such as quality control applications of herbal medicine. A study using a combination of UPLC–UV coupled with time-of-flight tandem mass spectrometry (RRLC–Q–TOF) method, reported a simultaneous qualitative and quantitative analysis of the main constituents present in various Niu Huang Jie Du Pill from different manufacturers (NHJDP, an ancient TCM, consists of 5 g of *Calculus Bovis Artificialis*, 50 g of *Realgar*, 200 g of *Gypsum Fibrosum*, 200 g of *Radix et Rhizoma Rhei Palmati*, 150 g of *Radix Scutellariae Baicalensis*, 100 g of *Radix Platycodi*, 25 g of *Borneol* and 50 g of *Radix Glycyrrhizae*). Ten major compounds could be detected in every sample: baicalin, baicalein, wogonoside, and wogonin from *Radix Scutellariae Baicalensis*; glycyrrhizic acid and liquiritin from *Radix Glycyrrhizae*; rhein, emodin, chrysophanol, and physcion from *Radix et Rhizoma Rhei Palmati*. Those compounds were reported to be responsible for the antipyretic and detoxicating effects of NHJDP. On a PCA score plot, they found that these compounds were responsible for the classification of different NHJDP preparations. The results were also compared with the conventional HPLC method where UPLC needed only 25 min to obtain similar separation, while conventional HPLC needed 80 min (Liang et al. 2010).

A metabolomics analysis of secondary metabolites of a pharmacologically active *Isatis tinctoria* leaves dichloromethane extract was carried out (Mohn et al. 2009). The activity of this medicinal plant is thought to be due to a synergistic interaction among its

compounds. Therefore, detailed information of the chemical composition of the plant is a prerequisite before addressing the possibility of pharmacological synergies. Metabolomics analysis was performed using multiple detection systems hyphenated with LC, including diode array detector (DAD), evaporative light scattering detector (Mangelsdorf et al. 1995), atmospheric pressure chemical ionization (APCI) and electrospray ionisation mass spectrometry (ESI–MS), and electrospray ionisation time-of-flight mass spectrometry (EI–TOF–MS) detectors. As a result, more than 65 compounds from diverse structural classes such as alkaloids, flavonoids, fatty acids, porphyrins, lignans, carotenoids, glucosinolates and cyclohexenones were identified.

UPLC coupled with photo-diode array detector and time-of-flight mass spectrometry (UPLC–PDA–TOF–MS) based chemical profiling approach was used to investigate the chemical equivalence of traditional and dispensing granule decoctions of traditional medicine combinatorial formulae (Li et al. 2010). A traditional granule decoction is made from the granule form of a single botanical, while a dispensing granule decoction is prepared by patients by mixing and dissolving different granules with hot water according to the recipe. There is a debate about the chemical components in dispensing granule decoction that might be different from those of a traditional decoction due to possible chemical interaction between mixed herbs. As a model, the authors chose San–Huang–Xie–Xin–Tang (SHXXT), a traditional combinatorial formula consisting of *Radix et Rhizoma Rhei* (*Rheum palmatum* L.), *Rhizoma Coptidis* (*Coptis chinensis* Franch) and *Radix Scutellariae* (*Scutellariae baicalensis* Georgie). SHXXT is commonly prescribed for treating hypertension and other various symptoms in oriental medicine. In this study, two different forms of SHXXT decoctions (traditional and dispensing granules) were subjected to UPLC–DAD–TOFMS analysis. The datasets of tR–m/z pairs, ion intensities and sample codes were analyzed with supervised orthogonal partial least squared discriminant analysis (OPLS–DA) to holistically compare the two kinds of samples. It was found that the preparations differed in the levels of berberine, epiberberine, palmatine, baicalin, wogonoside, and 2-O-galloyl-1-O-cinnamoylglucose which were higher in the dispensing granule decoction, and in the level of emodin which was higher in the traditional decoction (Li et al. 2010).

Also NMR-based metabolomics has been applied for quality control of botanicals. Kim et al. (2005) reported the analysis of *Ephedra* plant material by means of NMR. The four species that can grow in Taiwan can easily be distinguished and analyzing samples bought on the market can directly be recognized to the species level by using this method. Samples consisting of mixtures of species are immediately recognized. Moreover the ephedrine can be quantified by means of NMR (Kim et al. 2003; Kim et al. 2005).

Echinacea spp. are native North American plants widely used as herbal medicine and food supplement. Commercial *Echinacea* products often consist of substitutes of *Echinacea* plants which are not well defined chemically. Moreover, other factors may affect the chemical composition, and thus the efficacy of *Echinacea* preparations. The three most used medicinal *Echinacea* species, *E. purpurea*, *E. pallida*, and *E. angustifolia* grown under organic farming conditions, were subjected to supercritical fluid extraction (SFE) coupled with GC–MS to generate their metabolite profiles, which were compared using various approaches, including principal component analysis (PCA) and generalized association plots (GAP). Alkamides were identified as the major lipophilic constituents in *E. purpurea* and *E. angustifolia* root extracts and were found as the active anti-inflammatory compounds in this species. *Echinacea pallida* contains significantly less alkalamides, but higher amounts of ketoalkenes and ketoalkynes. The alkalamides content of the three *Echinacea* species was higher in roots than in aerial parts. The total alkalamides content of *E. purpurea* root was found to be significantly decreased by pre- and post-harvest abiotic stress, such as striking the leaves with a metal brush, simulated drought, incubation at 42 °C for 2 h, shade-drying, and sun-drying (Hou et al. 2010). In another study, by using UPLC–Q–TOF–MS based metabolomic, Lee et al. (2012) tried to differentiate *Aloe vera* L. leaves by size which reflects the leaves maturity stages. It was suggested that Aloin A, aloin B, and isoaloesin D were the major compounds responsible for the antioxidant activity of *Aloe vera* L. leaves since the compounds were found to be the highest in leaves which length was 20–30 cm. The leaves had the highest antioxidant activity among other samples with different size while the leaves commercially used in

industry are the ones with 50 cm length (Lee et al. 2012).

There are many other examples of the use of metabolomics for quality control of botanicals (van der Kooy et al. 2008), but none of them established a correlation between metabolites and activity. That means that the methods are only confirming identity but not quality in terms of the dose of the compounds related with activity.

Bioactivity—metabolite profiles correlation

The perfect quality control would be one that directly relates the activity to certain marker compounds in the metabolic profile. These compounds do not necessarily have an activity themselves, but could also be prodrugs or compounds that play a role in synergism. Several interesting examples have now been reported.

GC–MS-based metabolic profiling coupled with anti-inflammatory activity data of ginger (*Zingiber officinale*) from different sources and of some other *Zingiber* species showed that two other closely related *Zingiber* species had similar anti-inflammatory activity as *Z. officinale*, but there was no correlation with the content of gingerol, a compound specific to *Z. officinale*, and thought to be responsible for the anti-inflammatory activity. Therefore, it was concluded that other unidentified compounds contribute to this activity and need to be identified and quantified in order to guarantee the bioactivity of a particular *Zingiber* sample (Jiang et al. 2006).

The coupling of LC and TOF–MS was used to compare constituents and anticancer potential of gingers (*Zingiber officinale* Rosco, Zingiberaceae) treated in different ways (fresh, dried, and steamed at different temperature for different time: 100 °C for 1 h, or 120 °C for 0.5, 1, 2, 4, and 6 h). It was found that the steaming process at 120 °C for 4 h enhanced the anti-cancer effects of ginger on human HeLa cancer cells. Since the content of gingerols were decreased, while those of shogaols were increased with this treatment, the authors concluded that the shogaols in steamed ginger contribute to its anti-cancer potential (Cheng et al. 2011).

GC–EI–MS and LC–ESI–TOF–MS were used for metabolite profiling of *Pancreatium canariense* Ker Gawler (Amaryllidaceae). The alkaloids rich-

methanol extract of different parts of this plant showed medium acetylcholine esterase (AChE) inhibitory activity according to the order of bulbs (29.80 %), leaves (40.93 %), and fruits (58.06), respectively. Silylated methanolic extracts of bulbs, leaves and fruits of *P. canariense* were subjected to GC–MS analysis, where 102 compounds including some Amaryllidaceae alkaloids were detected. The concentrated bioactive alkaloid fractions were analyzed by GC–MS without any derivatisation step and as a result 31 alkaloids including galanthamine, haemanthamine, lycorine, homolycorine, tazettine and montanine were detected. Unknown compounds which had no reference spectra in the databases, were further subjected to LC–ESI–TOF–MS, through which additionally several other alkaloids were identified including a new compound, 3-*O*-acetylhabranthine (Torras-Claveria et al. 2010).

LC and LC–ESI–MS were used for fingerprinting of *Hypericum gentianoides* plant methanolic extracts. It was found that *H. gentianoides* has a very different chemical profile as compared to *H. perforatum* since it does not contain hypericin and hyperforin, instead it showed 9 metabolites with distinct UV absorption spectra in the active extract. Based on the UV spectra, retention times, and molecular ions of the compounds which were abundant in the active fraction, the authors presumed that the active compounds responsible for anti-inflammatory activity might be acyl-phloroglucinols (Hillwig et al. 2008).

Using targeted metabolomics data from conventional LC–DAD, PCA was able to differentiate Chenpi (dried peels of *Citrus tangerine* Tanaka), a Chinese traditional medicine for the treatment of indigestion and inflammatory syndromes), which was stored for 1 and 3 years, and heat treated (90 min at 120 °C, and 3 h at 120 °C). Total phenolic compounds and antioxidant activity were found to be higher in samples stored for longer period or had a longer heat treatment (Choi et al. 2011).

Rice bran has shown to have a beneficial effect on chronic diseases, and consequently it is used in functional food and dietary supplements. The results of a cell viability assay and a targeted approach using GC–MS, three rice varieties fermented with *Saccharomyces boulardii* were analyzed and the results were evaluated with multivariate data analysis (PCA, PLS-DA and OPLS). The results revealed a reduction in human B lymphomas by fermented rice bran extracts

from all three rice varieties as compared to all other non fermented controls. An altered metabolomic profile was found in the bioactive fermented varieties, where salicylic-, p-coumaric-, ferulic-, and caffeic acid, and also α -tocopherol and β -sitosterol were found as the major discriminating metabolites through PCA, PLS-DA and OPLS (significantly different as analyzed by student's *t* test and ANOVA) (Ryan et al. 2011).

The variation between plants is a disadvantage for the use, but can be an advantage in identifying the active compounds. Cardoso-Taketa et al. (2008) studied the sedative effect of the medicinal plant *Galphimia glauca* collected in 6 different places in Mexico. Out of the 6 accessions 2 were clearly much more active than the other 4. PCA analysis of the NMR spectra of all plants showed large differences between all plants, as one may expect, but no grouping of active and inactive plants. In such a case supervised multivariate data analysis methods such as PLS-DA and OPLS-DA can help by classifying all samples into two groups, one of actives and one of inactives. By doing so only the signals related to activity are considered and separated from all the “noise” of the large amount of variable compounds. Indeed a clear separation was observed in this approach and separation is based on triterpenoids that could be identified as the compounds that correlate with activity (Cardoso-Taketa et al. 2008).

Khatib and co-workers (2009) showed that PCA with TLC in combination with the results of an adenosine A1 receptor binding assay can lead to the identification of the active compounds in the leaves of *Orthosiphon stamineus* (Khatib et al. 2009).

As described above, one of the limitations in metabolomics is the polarity window of the extraction solvent. Active compounds might be missed because of low solubility. Moreover in complex extracts minor compounds might be active, but in screening on crude extracts these activities might be too low to be observed. Based on these considerations we have developed a novel method for extraction that is the comprehensive extraction. As previously described, this is a continuous extraction with a polarity gradient, yielding a number of fractions. After measuring the metabolome of each fraction by NMR they were subject to a receptor binding assay. By matching increased activity with increased level of a compound, the active compounds can be identified in the active fractions (Yuliana et al. 2011a, b). Similar technique

was used by Ching et al. (2012) to identify anti-platelet agent from the leaves of *Ardisia elliptica*, a Malay native plant traditionally used to treat chest pains. In this study, methanolic extract of the plant was partitioned with *n*-hexane, butanol, and water. The data of plant extracts effect on platelet aggregation in vitro was correlated to their chemical profile generated from GC–MS by using OPLS and PLS–DA. β -amyrin was one of the identified compounds with the highest correlation to the activity (Ching et al. 2012).

Proof of efficacy and identification mode of action

The major challenge in studies of medicinal plants is the identification of the active compounds and the target(s) they work on, i.e. the mode of action. A reductionist approach like bioassay guided fractionation can be used, but it is not always clear on what target one should focus, and working with in vivo tests usually quite large amounts are needed for a bioassay, which in such an approach is difficult to obtain. A systems biology approach in which bioassays are combined with metabolomics analyses is a novel and promising approach. Particularly when working with a whole organism, including clinical trials, one may see the total of the effects of a number of compounds present in the complex extracts, and even mixtures of extracts. That includes also synergism and prodrugs. This approach may lead to the discovery of novel bioactive compounds, but also in novel modes of action (and thus possibly novel targets). This new approach is now rapidly developing for studies on traditional medicine and is sometimes called reversed drug discovery or reversed pharmacology. Some examples may illustrate the enormous potential of this approach.

The use of metabolomic profiling to study the mode of action of active natural products recently first emerged in antimicrobial activity analysis. For example, HPLC–DAD–ESI–MS application followed by PCA confirmed that dihydrocucurbitacin F-25-*O*-acetate was the main component having an antimicrobial activity in *Hemsleya pengxianensis*. The possible antibacterial mode of action on *Staphylococcus aureus* CCTCC AB9105 was predicted by comparing the metabolite profile of *S. aureus* treated by the plant extract and commercial antibacterials with known mode of action. It was found that the inhibition of cell wall synthesis is the possible antibacterial mode of

dihydrocucurbitacin F-25-*O*-acetate and *H. pengxianensis* (Biao-Yi et al. 2008).

Epimedium brevicornum Maxim, one of the most popular herbal medicines in China is used for toning the kidney and strengthening bones, but the kind of constituents responsible for its pharmacological effect, and the mechanism through which it affects the entire body metabolism was unknown. The effect of *E. brevicornum* on the restoration of kidney abnormalities of rats treated with a high dose of hydrocortisone was reported. The injection of hydrocortisone at a high dose induces a pathological condition known as kidney deficiency in TCM. The plasma metabolites of rats from three different conditions (pre-hydrocortisone intervention, post-hydrocortisone intervention, and *E. brevicornum* Maxim. treated) were analyzed using UPLC–MS followed by PCA. A distinct increase of the level of three metabolites was found after hydrocortisone intervention, one of them was identified as ethylindole-3-crylic acetate. After treatment with *E. brevicornum*, the plasma metabolite profile was close to that of the pre-hydrocortisone intervention. Of the four compounds of this herbal preparation detected in plasma (epimedin C, icariin, icaraside II, and 2''-*O*-rhamnosoyl icaride), epimedin C and icariin were also found in urine, and it was concluded that those might be related to the activity (Famei Li et al. 2007).

The efficacy and the mechanism of action of the traditional Chinese medicine preparation known as Xindi soft capsules (consisting of sea buckthorn flavonoids and sea buckthorn berry oil) used to treat blood stasis, were studied by UPLC–QTOF–MS followed by PCA and PLS–DA analysis of the data. Blood stasis includes hematological disorders such as haemorrhage, nasal congestion, thrombosis, and local ischemia (microclots). The comparison of the urinary metabolite pattern obtained for 5 groups of rats (healthy control group, acute blood stasis model group, low dose group of Xindi soft capsule, middle dose group of Xindi soft capsule, high dose group of Xindi soft capsule), showed that the treated groups were located between the acute blood stasis model group and the healthy control group. Confirmation by hemorheological analysis proved that within the treated group, especially in the case of the highest dose group, the urine metabolite pattern tends toward that of the control group. Some potential biomarkers were identified as well, such as cholic acid, phenylalanine and kynurenic acid (Zhao et al. 2008).

Dai and co-workers (2011) used UFLC–MS–IT–TOF to study the preventive efficacy and the metabolic changes caused by the traditional Chinese medicine tongxinluo (TX) in endothelial dysfunction rats. Simvastatin (SV), a lipid lowering drug from statins class was used as a control. Initially, based on PCA results, the authors identified several markers which differentiating urine profile of healthy rats with endothelial dysfunction rats. Further analysis using PCA revealed that the urine profile of endothelial dysfunction rats treated with TX was more similar to the healthy rats than those of endothelial dysfunction rats treated with SV. The study also showed that TX exhibited a preventive action against endothelial dysfunction by regulating multiple metabolic pathways to their normal state, whereas simvastatin affected only selected pathways (Dai et al. 2011).

Suanzaoren decoction (SZRD) is an ancient TCM formula for sleeping disorder. SZRD consists of five herbal medicines as follow: *Ziziphus jujuba* Mill var. *spinosa* (Bunge) Huex. H.F. Chou. seeds, *Ligusticum chuanxiong* Hort. rhizomes, *Anemarrhena asphodeloides* Bge. rhizomes, and *Glycyrrhiza uralensis* Fisch. Roots. The effect of SZRD intervention in an insomnia *Drosophila* model was studied by using UPLC–ESI–SYNAPT–HDMS combined with pattern recognition approaches including PCA, PLS–DA and OPLS–DA (Yang et al. 2012). Initially they characterized 9 insomnia biomarkers in negative mode and 11 in positive mode MS detection. In the PCA score plot, the treated *Drosophilas* were shifted back close to the normal ones as compared to the non-treated ones. Since they found that serotonin, melatonin, 5-hydroxy-L-tryptophan, and prostaglandin D2 have the strongest association with normal characteristics, they concluded that SZRD acts through serotonergic activation.

An HPLC–DAD–ESI–MS technique coupled with PCA was used to study the synergistic interaction between *Salvia miltiorrhiza* and *Dalbergiae odoriferae*. Healthy rabbits and rabbits with Qi-stagnancy and blood stasis were the objects of the study. They were fed by *S. miltiorrhiza* or a mixture of *S. miltiorrhiza* with *D. odoriferae* water extract. The blood plasma was collected at several time points after extract administration and further subjected to chromatography analysis. The PCA results revealed that Qi-stagnancy and blood stasis were strongly correlated with the presence of 4-methylbenzamide in the plasma, and

that synergism between *S. miltiorrhiza* with *D. odoriferae* was strongly associated with biomarker β -(3,4-dihydroxyphenyl)- α -hydroxy acid isopropyl ester (Wang et al. 2011).

The anti-diabetic mechanism of berberine, a quaternary alkaloid, was studied by Yan Gu and co-workers (2010). UPLC–QTOF MS was used to identify metabolites of blood plasma of human subjects treated with the compounds. There was a significant decrease of 10 fatty acids in berberine treated patients as compared to placebo, which may explain the mechanism of berberine in diabetes treatment, i.e. by down-regulating the high level of serum free fatty acids which means further mediating lipids and glucose metabolism (Gu et al. 2010).

Dai et al. (2010) reported the use of GC–MS to study the effect of a famous Chinese prescription, Xiaoyaosan, in rats. This herbal is traditionally used for depression and mental disorder treatment and composed of *Poria* (*Poria cocos* (Schw.)Wolf), *Radix Paeoniae Alba* (*Paeonia lactiflora* Pall.), *Radix Glycyrrhizae* (*Glycyrrhiza uralensis* Fisch.), *Radix Bupleuri* (*Bupleurum chinense* DC.), *Radix Angelicae Sinensis* (*Angelica sinensis* (Oliv.) Diels), *Rhizoma Atractylodis Macrocephalae* (*Atractylodes macrocephala* Koidz.), *Herba Menthae* (*Mentha haplocalyx* Briq.), and *Rhizoma Zingiberis Recens* (*Zingiber officinale* Rosc.). The urine profiles of rats treated with this herbal at three doses (low, medium, and high), together with control groups (no stress and no herbal treatment, stress without herbal treatment), and a group with stress and amitriptyline (reference drug) at low dose treatment, were measured by GC–MS and then the plotted to PCA and PLS–DA. On the score plot there was a clear time-related trajectory of metabolite patterns at different time points (day 1, 7, 12, 14, 18, and 21) where metabolites changes of a high dose group of xiaoyaosan showed recovering tendency to normal level at the end of dosing period. Thirteen urine metabolites were elucidated from the loading plot as potential markers for the anti-depression effect of Xiaoyaosan (Dai et al. 2010).

Conclusion

The paradigm of drug discovery is “single drug single target”, but with the sharp decrease in novel low molecular weight pharmaceuticals, there is a trend

towards novel approaches to find new leads for drug development. Natural products are an important source for drug development, but in the traditional way of bioassay guided fractionation, this approach is too elaborate and costly to be successful in an industrial setting. As traditional medicine is a pre-selected group of materials with an increased chance of finding novel leads, novel modes of action and novel targets, there is a clear increasing interest in studying this heritage. On the one hand to confirm the traditional use and thus make them into evidence-based botanicals, on the other hand to find through reversed drug discovery novel leads. The ‘omics’ can play an important role in such studies. As illustrated with a series of examples metabolomics can be used on the side of the botanical (different accessions, extracts or fractions) to identify by correlating with the activity what metabolites are involved in the activity. This can be used for quality control as well as for further in depth studies of the mode of action. The mode of action can be learned from metabolomic analysis of the test system. Such a systems biology approach to studies of the 30,000–70,000 medicinal plants holds a lot of promise!

References

- Badjakov I, Nikolova M, Gevrenova R, Kondakova V, Todorovska E, Atanassov A (2008) Bioactive compounds in small fruits and their influence on human health. *Bio-technol Biotechnol Eq* 22:581–587
- Biao-Yi Z, Yu Y, Zeng-Liang Y (2008) Investigation of antimicrobial model of *Hemsleya pengxianensis* W.J. Chang and its main active component by metabolomics technique. *J Ethnopharmacol* 116:89–95
- Cardoso-Taketa AT, Pereda-Miranda R, Choi YH, Verpoorte R, Villareal ML (2008) Metabolic profiling of the Mexican anxiolytic and sedative plant *Galphimia glauca* using Nuclear Magnetic Resonance spectroscopy and Multivariate Data Analysis. *Planta Med* 74:1295–1301
- Chang W-T, Choi YH, Van der Heijden R, Lee M-S, Lin M-K, Kong H, Kim HK, Verpoorte R, Hankemeier T, Van der Greef J, Wang M (2011) Traditional processing strongly affects metabolite composition by hydrolysis in *Rehmannia glutinosa* roots. *Chem Pharm Bull* 59:546–552
- Cheng X-L, Liu Q, Peng Y-B, Qi L-W, Li P (2011) Steamed ginger (*Zingiber officinale*): changed chemical profile and increased anticancer potential. *Food Chem* 129:1785–1792
- Ching J, Soh W-L, Tan C-H, Lee J-F, Tan J-YC, Yang J, Yap C-W, Koh H-L (2012) Identification of active compounds from medicinal plant extracts using gas chromatography-mass spectrometry and multivariate data analysis. *J Sep Sci* 35:53–59
- Choi M-Y, Chai C, Park JH, Lim J, Lee J, Kwon SW (2011) Effects of storage period and heat treatment on phenolic compound composition in dried Citrus peels (Chenpi) and discrimination of Chenpi with different storage periods through targeted metabolomic study using HPLC–DAD analysis. *J Pharm Biomed Anal* 54:638–645
- Colquhoun IJ (2007) Use of NMR for metabolic profiling in plant systems. *J Pestic Sci* 32:200–212
- Dai Y, Li Z, Xue L, Dou C, Zhou Y, Zhang L, Qin X (2010) Metabolomics study on the anti-depression effect of xiaoyaosan on rat model of chronic unpredictable mild stress. *J Ethnopharmacol* 128:482–489
- Dai W, Wei C, Kong H, Jia Z, Han J, Zhang F, Wu Z, Gu Y, Chen S, Gu Q, Lu X, Wu Y, Xu G (2011) Effect of the traditional Chinese medicine tongxinluo on endothelial dysfunction rats studied by using urinary metabolomics based on liquid chromatography-mass spectrometry. *J Pharm Biomed Anal* 56:86–92
- Feher M, Schmidt JM (2003) Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J Chem Inf Comput Sci* 43:218–227
- Gu Y, Zhang Y, Shi X, Li X, Hong J, Chen J, Gu W, Lu X, Xu G, Ning G (2010) Effect of traditional Chinese medicine berberine on type 2 diabetes based on comprehensive metabolomics. *Talanta* 81:766–772
- Han J, Datla R, Chan S, Borchers CH (2009) Mass spectrometry-based technologies for high-throughput metabolomics. *Bioanalysis* 1:1665–1684
- Henkel T, Brunne RM, Müller H, Reichel F (1999) Statistical investigation into the structural complementarity of natural products and synthetic compounds. *Angew Chem Int Ed* 38:643–647
- Hillwig ML, Hammer KDP, Birt DF, Wurtele ES (2008) Characterizing the metabolic fingerprint and anti-inflammatory activity of *Hypericum gentianoides*. *J Agric Food Chem* 56:4359–4366
- Hou CC, Chen CH, Yang NS, Chen YP, Lo CP, Wang SY, Tien YJ, Tsai PW, Shyr LF (2010) Comparative metabolomics approach coupled with cell- and gene-based assays for species classification and anti-inflammatory bioactivity validation of *Echinacea* plants. *J Nutr Biochem* 21:1045–1059
- Ingkaninan K, Hazekamp A, de Best CM, Irth H, Tjaden UR, van der Heijden R, van der Greef J, Verpoorte R (2000) The application of HPLC with on-line coupled UV/MS: biochemical detection for isolation of an acetylcholinesterase inhibitor from *Narcissus* Sir Winston Churchill. *J Nat Prod* 63:803–806
- Jiang H, Xie Z, Koo HJ, McLaughlin SP, Timmermann BN, Gang DR (2006) Metabolic profiling and phylogenetic analysis of medicinal *Zingiber* species: tools for authentication of ginger (*Zingiber officinale* Rosc.). *Phytochemistry* 67:1673–1685
- Khatib A, Yuliana N, Jinap S, Sarker M, Jaswir I, Wilson E, Chung S, Verpoorte R (2009) Identification of possible compounds possessing adenosine A1 receptor binding activity in the Leaves of *Orthosiphon stamineus* using TLC

- and multivariate data analysis. *J Liq Chrom Rel Technol* 32:2906–2916
- Kim HK, Verpoorte R (2010) Sample preparation for plant metabolomics. *Phytochem Anal* 21:4–13
- Kim HK, Choi YH, Chang W-T, Verpoorte R (2003) Quantitative analysis of ephedrine analogues from *Ephedra* species using ¹H-NMR. *Chem Pharm Bull* 51:1382–1385
- Kim HK, Choi YH, Erkelens C, Lefeber AWM, Verpoorte R (2005) Metabolic fingerprinting of *Ephedra* species using ¹H-NMR spectroscopy and Principal Component Analysis. *Chem Pharm Bull* 53:105–109
- Kim HK, Choi YH, Verpoorte R (2010) NMR-based metabolic analysis of plants. *Nat Protocols* 5:536–549
- Koehn FE (2008) High impact technologies for natural products screening. In: Amstutz R, Petersen F (eds) *Natural compounds as drugs volume I*. Birkhäuser, Basel, pp 175–210
- Lee S, Do S-G, Kim SY, Kim J, Jin Y, Lee CH (2012) Mass spectrometry-based metabolite profiling and antioxidant activity of *Aloe vera* (*Aloe barbadensis* Miller) in different growth stages. *J Agric Food Chem* 60:11222–11228
- Li F, Lu X, Liu H, Liu M, Xiong Z (2007) A pharmacometabonomic study on the therapeutic basis and metabolic effects of *Epimedium brevicornum* Maxim on hydrocortisone-induced rat using UPLC-MS. *Biomed Chromatogr* 21:397–405
- Li S-L, Song J-Z, Qiao C-F, Zhou Y, Xu H-X (2010) UPLC-PDA-TOFMS based chemical profiling approach to rapidly evaluate chemical consistency between traditional and dispensing granule decoctions of traditional medicine combinatorial formulae. *J Pharm Biomed Anal* 52:468–478
- Liang X, Zhang L, Zhang X, Dai W, Li H, Hu L, Liu H, Su J, Zhang W (2010) Qualitative and quantitative analysis of traditional Chinese medicine Niu Huang Jie Du Pill using ultra performance liquid chromatography coupled with tunable UV detector and rapid resolution liquid chromatography coupled with time-of-flight tandem mass spectrometry. *J Pharm Biomed Anal* 51:565–571
- Ma C, Wang H, Lu X, Xu G, Liu B (2008) Metabolic fingerprinting investigation of *Artemisia annua* L. in different stages of development by gas chromatography and gas chromatography-mass spectrometry. *J Chromatogr A* 1186:412–419
- Maltese F, van der Kooy F, Verpoorte R (2009) Solvent derived artifacts in natural products chemistry. *Nat Prod Commun* 4:447–454
- Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schütz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P, Evans RM (1995) The nuclear receptor superfamily: the second decade. *Cell* 83:835–839
- Mohn T, Plitzko I, Hamburger M (2009) A comprehensive metabolite profiling of *Isatis tinctoria* leaf extracts. *Phytochemistry* 70:924–934
- Newman DJ, Cragg GM (2007) Natural products as sources of new drugs over the last 25 years. *J Nat Prod* 70:461–477
- Newman DJ, Cragg GM (2012) Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod* 75:311–335
- Newman DJ, Cragg GM, Snader KM (2003) Natural products as sources of new drugs over the period 1981–2002. *J Nat Prod* 66:1022–1037
- Rhee IK, van de Meent M, Ingkaninan K, Verpoorte R (2001) Screening for acetylcholinesterase inhibitors from Amaryllidaceae using silica gel thin-layer chromatography in combination with bioactivity staining. *J Chromatogr A* 915:217–223
- Ryan EP, Heuberger AL, Weir TL, Barnett B, Broeckling CD, Prenni JE (2011) Rice bran fermented with *Saccharomyces boulardii* generates novel metabolite profiles with bioactivity. *J Agric Food Chem* 59:1862–1870
- Schenk T, Breel GJ, Koevoets P, van Den Berg S, Hogenboom AC, Irth H, Tjaden UR, van Der Greef J (2003) Screening of natural products extracts for the presence of phosphodiesterase inhibitors using liquid chromatography coupled online to parallel biochemical detection and chemical characterization. *J Biomol Screen* 8:421–429
- Schobel U, Frenay M, Van Elswijk DA, McAndrews JM, Long KR, Olson LM, Bobzin SC, Irth H (2001) High resolution screening of plant natural product extracts for estrogen receptor α and β binding activity using an online HPLC-MS biochemical detection system. *J Biomol Screen* 6:291–303
- Shyur L-F, Yang N-S (2008) Metabolomics for phytomedicine research and drug development. *Curr Opin Chem Biol* 12:66–71
- Steinmann D, Ganzera M (2011) Recent advances on HPLC/MS in medicinal plant analysis. *J Pharm Biomed Anal* 55:744–757
- Torras-Claveria L, Berkov S, Jáuregui O, Caujapé J, Viladomat F, Codina C, Bastida J (2010) Metabolic profiling of bioactive *Pancreatium canariense* extracts by GC-MS. *Phytochem Anal* 21:80–88
- van der Kooy F, Verpoorte R, Marion Meyer JJ (2008) Metabolomic quality control of claimed anti-malarial *Artemisia afra* herbal remedy and *A. afra* and *A. annua* plant extracts. *S Afr J Bot* 74:186–189
- van der Kooy F, Maltese F, Choi Y (2009) Quality control of herbal material and phytopharmaceuticals with MS and NMR based metabolic fingerprinting. *Planta Med* 75:763–775
- van Elswijk D, Schobel U, Lansky E, Irth H, van der Greef J (2004) Rapid dereplication of estrogenic compounds in pomegranate (*Punica granatum*) using on-line biochemical detection coupled to mass spectrometry. *Phytochemistry* 65:233–241
- Verpoorte R, Choi Y, Kim H (2007) NMR-based metabolomics at work in phytochemistry. *Phytochem Rev* 6:3–14
- Verpoorte R, Choi Y, Mustafa N, Kim H (2008) Metabolomics: back to basics. *Phytochem Rev* 7:525–537
- Wang M, Lamers R-JAN, Korthout HAAJ, van Nesselrooij JHH, Witkamp RF, van der Heijden R, Voshol PJ, Havekes LM, Verpoorte R, van der Greef J (2005) Metabolomics in the context of systems biology: bridging traditional Chinese medicine and molecular pharmacology. *Phytother Res* 19:173–182
- Wang S-X, Luo K, Liang J, Fan F, Li H, Zheng J-B, Zheng X-H (2011) Metabolomics study on the synergistic interaction between *Salvia miltiorrhiza* and *Lignum dalbergiae odoriferae* used as “Jun-Shi” herbs in a *S. miltiorrhiza* recipe. *Med Chem Res* 20:16–22
- Yang B, Zhang A, Sun H, Dong W, Yan G, Li T, Wang X (2012) Metabolomic study of insomnia and intervention effects of

- Suanzaoren decoction using ultra-performance liquid-chromatography/electrospray-ionization synapt high-definition mass spectrometry. *J Pharm Biomed Anal* 58:113–124
- Yi L, Yuan D, Liang Y, Xie P, Zhao Y (2007) Quality control and discrimination of *Pericarpium Citri Reticulatae* and *Pericarpium Citri Reticulatae Viride* based on high-performance liquid chromatographic fingerprints and multivariate statistical analysis. *Anal Chim Acta* 588:207–215
- Yuliana ND, Khatib A, Choi YH, Verpoorte R (2011a) Metabolomics for bioactivity assessment of natural products. *Phytother Res* 25:157–169
- Yuliana ND, Khatib A, Verpoorte R, Choi YH (2011b) Comprehensive extraction method integrated with NMR metabolomics: a new bioactivity screening method for plants, adenosine A1 receptor binding compounds in *Orthosiphon stamineus* Benth. *Anal Chem* 83:6902–6906
- Zhao X, Zhang Y, Meng X, Yin P, Deng C, Chen J, Wang Z, Xu G (2008) Effect of a traditional Chinese medicine preparation Xindi soft capsule on rat model of acute blood stasis: a urinary metabolomics study based on liquid chromatography-mass spectrometry. *J Chromatogr B* 873:151–158