



**HAL**  
open science

## Metabolomics for Bioactivity Assessment of Natural Products

Nancy Dewi Yuliana, Alfi Khatib, Young Hae Choi, Robert Verpoorte

► **To cite this version:**

Nancy Dewi Yuliana, Alfi Khatib, Young Hae Choi, Robert Verpoorte. Metabolomics for Bioactivity Assessment of Natural Products. *Phytotherapy Research*, 2010, 10.1002/ptr.3258 . hal-00599835

**HAL Id: hal-00599835**

**<https://hal.science/hal-00599835>**

Submitted on 11 Jun 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Metabolomics for Bioactivity Assessment of Natural Products**

Journal:	<i>Phytotherapy Research</i>
Manuscript ID:	PTR-10-0117.R2
Wiley - Manuscript type:	Review
Date Submitted by the Author:	30-May-2010
Complete List of Authors:	Dewi Yuliana, Nancy; Institue of Biology, Leiden University Khatib, Alfi; Bogor Agricultural University Choi, Young Hae; Institute of Biology, Leiden University Verpoorte, Robert; Institue of Biology, Leiden University
Keyword:	natural product, drug discovery, metabolomics, bioactivity



view

## Metabolomics for Bioactivity Assessment of Natural Products

Nancy Dewi Yuliana<sup>1,2</sup>, Alfi Khatib<sup>2,3</sup>, Young Hae Choi<sup>1</sup>, Robert Verpoorte<sup>1</sup>

<sup>1</sup> Division of Pharmacognosy, Section of Metabolomics, Institute of Biology, Leiden University, Leiden, The Netherlands

<sup>2</sup> Department of Food Science and Technology, Bogor Agricultural University, Bogor, Indonesia

<sup>3</sup> Center of Excellence for Food Safety Research, Faculty of Food Science and Technology, University Putra Malaysia, 43400 Serdang, Selangor Darul Ehsan, Malaysia.

Natural products have historically been a rich source of lead molecules in drug discovery based on their capability to create unique and diverse chemical structures. However, it is also true that the vast number of metabolites typically present in natural products and their huge dynamic range results in the loss of many possibly bioactive natural compounds, becoming an inextricable obstacle for drug development. Recently, new strategies which favor a holistic approach as opposed to the traditional reductionist methods previously used, have been introduced with the purpose of overcoming the bottlenecks in natural product research. This approach is based on the application of new technologies, including metabolomics, for example. Metabolomics allows a systematic study of a complex mixture such as phytochemical preparation, which can be linked to observations obtained through biological testing systems without the need of isolating active principles. This may put drug discovery from natural products back in the limelight again. In this review paper, the description of some examples of successful metabolomics applications in several important fields related to drug discovery from natural sources aims at raising the potential of metabolomics in reducing the gap between natural products (NP) and modern drug discovery demand.

1  
2  
3  
4  
5  
6  
7 **Key words:** natural product, drug discovery, bioactivity, metabolomics.  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For Peer Review

## Introduction

In recent years the development of novel drugs has become increasingly difficult, expensive and time consuming. It is estimated that more than 800 million USD and an average of 14.2 years are needed for a novel drug application (NDA) approval. Both factors are rapidly increasing with time while the output of newly launched drugs has dropped (Butler, 2004; Dickson and Gagnon, 2004). The introduction of high throughput screening (HTS), a process which allows the testing of a large number of samples on certain molecular targets, was expected to shorten the discovery route. To fully use its high potential, HTS requires a large number of compounds to screen, in the order of thousands of samples per day (Mishra *et al.*, 2008). In order to meet this demand, high throughput synthesis and combinatorial chemistry were developed. Concomitantly, further development in the field of structural biology and genomics, e.g. cloning and expression technology resulting in the production of various purified human receptors and enzymes which could be studied at a molecular level, produced a sharp increase in the number of possible molecular targets, demanding more drug-based libraries of a wide chemical diversity to screen (Koehn and Carter, 2005). The combination of these new strategies seemed very powerful and the delivery of a large number of novel leads was expected (Ortholand and Ganesan, 2004). However, these expectations were shown to be greatly overrated as the number of new active compounds that made it to the market in the past two decades reached its lowest level ever in 2002 with an insignificant improvement in 2003 (Newman and Cragg, 2007; Newman *et al.*, 2003; Ortholand and Ganesan, 2004). Sorafenib, a multikinase inhibitor indicated for advanced renal cancer, is the single synthetic drug derived from combinatorial chemistry approved by Food and Drug Administration (FDA) for clinical

1  
2  
3  
4 use (Newman and Cragg, 2007). The fact, this increased HTS screening capacity and  
5  
6 number of novel molecular targets has failed to produce an increase of new lead  
7  
8 compounds poses the question on whether this is due to limitations in the innovative  
9  
10 aspects of the discovery process itself, e.g. generation of compounds limited chemical  
11  
12 diversity, or because of the potential of natural products (NP) has been de-emphasized  
13  
14 in the past few years (Butler, 2004).  
15  
16  
17

18  
19 The possibility of combining NP extract libraries with HTS also seems interesting,  
20  
21 though not always an easy task. Several reasons for the incompatibility of NP and HTS  
22  
23 have been described (Butler, 2004; Ortholand and Ganesan, 2004). The complexity of  
24  
25 NP extract which is a complex mixture of mostly uncharacterized compounds which  
26  
27 may mask the bioactivity of target compounds or cause false positive results is the most  
28  
29 common hurdle. The large dynamic range (e.g. the difference in the levels of major  
30  
31 compounds as compared to that of minor compounds, differences in polarity, in boiling  
32  
33 or evaporating points, etc) which may affect compounds solubility makes the integration  
34  
35 of NP and HTS more complicated. The differences between areas where plants are  
36  
37 grown, the methods of production and processing of plants are also important since  
38  
39 these might affect the chemical profile of the NP (Jahangir *et al.*, 2008; Wang *et al.*,  
40  
41 2005). The intellectual property issue in natural product outsourcing is another problem.  
42  
43  
44  
45 Thus, the drug discovery process using plant sources was almost abandoned (Dickson  
46  
47 and Gagnon, 2004; McChesney *et al.*, 2007). Companies such as Glaxo Wellcome,  
48  
49 Smith Kline & French, and Pfizer, gradually eliminated their NP screening programmes  
50  
51 while others such as Merck and Novartis continued their exploration but focussed on  
52  
53 microbial sources (Ortholand and Ganesan, 2004).  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5 The failure to increase the number of new drugs delivered to the market has placed  
6  
7 the drugs discovery programs of pharmaceutical companies under scrutiny. Looking  
8  
9 back to the past two decades, of the 877 novel medicines developed between 1981 and  
10  
11 2002, 6% were NP, 27% were NP derivatives and 16% were synthetic products  
12  
13 developed on the basis of a NP model (Newman *et al.*, 2003), indicating that NP is still  
14  
15 an important source of novel leads for therapeutical drugs. This situation has led to a  
16  
17 renewed interest in NP in the drug discovery programs. Therefore, it is a challenging  
18  
19 task to make NP available for the drug discovery arena.  
20  
21  
22

23  
24 Most NP drug discovery work follows ‘the silver bullet’ approach as is usual in  
25  
26 present Western medicine, in which the search is focused on single compounds with a  
27  
28 particular biological activity (single target-single compound paradigm) tested at  
29  
30 molecular level (e.g. receptor binding inhibition or enzyme screening) (Verpoorte *et al.*,  
31  
32 2005). This approach which has been responsible for the successful delivery of many  
33  
34 potent drugs in the past, such as the statins to lower cholesterol level and huperzine A  
35  
36 for Alzheimer’s disease (Li *et al.*, 2008b), is not actually the best approach for NP  
37  
38 studies since the presence of synergism and pro-drugs in an extract can be misleading  
39  
40 (Verpoorte *et al.*, 2005). It is also one of the reasons for the long time line and high  
41  
42 costs which the pharmaceutical companies have to deal with when working with NP.  
43  
44 Shifting the silver bullet paradigm to a more holistic approach, which is actually the  
45  
46 philosophy of some traditional medicine such as Traditional Chinese Medicine (TCM)  
47  
48 and Ayurveda, is now evolving into a new trend (Ulrich-Merzenich *et al.*, 2007;  
49  
50 Verpoorte *et al.*, 2005; Wang *et al.*, 2005). In a holistic approach, a whole herbal  
51  
52 preparation, which consists of numerous multi components, is viewed as one active  
53  
54 entity with effects on multiple targets in living organisms (Li *et al.*, 2008b; Ulrich-  
55  
56  
57  
58  
59  
60

1  
2  
3  
4 Merzenich *et al.*, 2007). For such an approach, the response of the whole system needs  
5  
6 to be observed with all possible tools such as clinical trials or experimental animal  
7  
8 models, where both classic physiological observations (e.g. blood pressure, analgesic  
9  
10 activity, sedative) and modern molecular observation (gene expression, proteome,  
11  
12 metabolome) can be used to obtain better insight into different changes in organism.  
13  
14 This is known as a systems biology approach (Verpoorte *et al.*, 2005).  
15  
16  
17

18  
19 Among these ‘omics’ methods, metabolomics, a relatively new member, is thought  
20  
21 to be the most informative as regards understanding systems biology since it reflects  
22  
23 the genotype (Sumner *et al.*, 2003). Therefore it is a valuable tool to use in the provision  
24  
25 of the information required in this new approach which is particularly interesting for the  
26  
27 study of medicinal plants. It thus offers new opportunities for NP in the drug discovery  
28  
29 arena.  
30  
31  
32

33  
34 This review will discuss the chemical diversity of compounds present in NP as  
35  
36 compared to synthetic compounds, the obstacles which might limit NP integration with  
37  
38 the new emerging drug discovery technologies, the shift in tendency from a reductionist  
39  
40 to a holistic approach in NP research and how metabolomics can help to narrow the gap  
41  
42 between NP and the demand of modern drug discovery, and helping to turn NP and  
43  
44 medicinal plants into a unique drug source.  
45  
46  
47  
48

### 49 **1. Natural products as a drug source; their advantages and limitations**

50  
51 Factors which distinguish NP from synthetic chemicals include their incomparable  
52  
53 molecular diversity and biological functionality (Nisbet and Moore, 1997). Starting with  
54  
55 only a few building blocks of NP provided by the biosynthetic pathways, the action of  
56  
57 various “decorating” enzymes provide a large diversity of end products with different  
58  
59  
60



1  
2  
3  
4 functionalities (Nisbet and Moore, 1997; Pieters and Vlietinck, 2005; Verpoorte, 1998).  
5  
6 Compared to products resulting from combinatorial chemistry libraries which usually  
7  
8 lack chirality and structure rigidity, two important requirements for pharmaceutical  
9  
10 drug-like properties (Nisbet and Moore, 1997), NP have privileged structures which are  
11  
12 able to interact with proteins and other biological targets for specific purposes (Koehn  
13  
14 and Carter, 2005).  
15  
16  
17

18  
19 To address the question of how the unique structural properties of NP are able to  
20  
21 complement the chemical space of synthetic compounds, the comparison of the major  
22  
23 structural differences between combinatorial compounds, drugs, and NP molecules has  
24  
25 been made. It is obvious that current commercial drugs and NP share more similarities  
26  
27 in several topological pharmacophore patterns. The important differences between NP  
28  
29 and synthetic compounds include the number of chiral centers, the prevalence of  
30  
31 aromatic rings, the presence of complex ring systems, and the degree of the saturation of  
32  
33 the molecule as well as the number and ratios of different heteroatoms (Feher and  
34  
35 Schmidt, 2003). Natural products have a higher number of chiral centers, a higher  
36  
37 number of oxygen atoms instead of nitrogen, sulfur and halogen containing groups  
38  
39 present than in synthetic compounds. Natural products have a lower ratio of aromatic  
40  
41 ring atoms to total heavy atoms, and a higher number of solvated hydrogen bond donors  
42  
43 and acceptor functional groups. These characteristics are favorable attributes for a drug  
44  
45 candidate profile (Feher and Schmidt, 2003; Henkel *et al.*, 1999; Newman and Cragg,  
46  
47 2007; Newman *et al.*, 2003). With these unique properties, in general, NP actually have  
48  
49 more drug-like properties than synthetic compounds (Feher and Schmidt, 2003; Koehn,  
50  
51 2008).  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

However, a NP-based drug discovery project also poses some challenges, mostly connected with their presence within a complex matrix with all kind of active compounds (Bindseil *et al.*, 2001; Koehn, 2008; Koehn and Carter, 2005; McChesney *et al.*, 2007; Ortholand and Ganesan, 2004; Pieters and Vlietinck, 2005; Schmid *et al.*, 1999):

(1) The first challenge is dereplication which is the rapid identification of known and unknown compounds from partially purified mixtures. In a number of test systems, common plant products which have been found to be active e.g. unsaturated fatty acids on adenosine and GABA receptors or tannins on enzymes based bioassays, have to be identified at the very beginning of the study of an active extract.

(2) Interaction between metabolites present in extracts. Extracts consist of mixtures of many compounds requiring an elaborate purification to isolate new pure compounds. Another problem of complex extracts is the presence of antagonism or synergisms which might occur between components in a mixture, posing a major intellectual challenge which can be painstakingly time consuming.

(3) Structure determination by spectroscopic methods.

(4) Large scale production of the active component is sometimes an obstacle as the identified NP lead might be a minor compound of the extract, or the supply of the material for extraction might be limited.

(5) Sourcing of authenticated plants, a very important issue for quality control. For example, some plants have congener substitutes which have an active principal content which may be different to the original one.

1  
2  
3  
4 Last but not least, biodiversity and intellectual property issues can be another  
5  
6 complicating factor related to this context.  
7  
8  
9

## 10 11 **2. Previous efforts to increase the number of compounds for drug candidates**

12 Because of the demand to speed up the drug discovery timeline and to reduce costs, the  
13  
14 *in vitro* screening system has become the preferred method. The use of animal tests or  
15  
16 isolated organs was abandoned for more specific testing targets at a molecular level  
17  
18 (reductionist mode) (Verpoorte *et al.*, 2006). Following this trend, there are two crucial  
19  
20 prerequisites to start the drug discovery process: a large high quality library of  
21  
22 candidates to be screened, and a reliable and sensitive assay to screen the activity.  
23  
24 Combinatorial chemistry was developed to address the first need (Williard *et al.*, 1996).  
25  
26 Based on the total chemical space, this technology is theoretically able to provide  $10^{30}$  –  
27  
28  $10^{100}$  compounds (Macarron, 2006; Myers, 1997).  
29  
30  
31  
32  
33  
34

35 The second need was met by high throughput screening (HTS) as the response to the  
36  
37 availability of hundreds to thousands of compounds resulting from combinatorial  
38  
39 chemistry. HTS is a fast and cost-effective technology designed to select drug  
40  
41 candidates (Bleicher *et al.*, 2003; Gómez-Hens and Aguilar-Caballo, 2007). It is  
42  
43 defined as an automatic approach in drug discovery which is characterized by  
44  
45 miniaturization to allow the use of minute quantities of samples and reagent.  
46  
47 Continuous innovations keep taking place to maximize the quantity of screened  
48  
49 compounds and to minimize the cost. The miniaturization of screening assay aimed at  
50  
51 increasing HTS capacity is achieved by increasing the screening plate density and  
52  
53 minimizing screening volume (Landro *et al.*, 2000). The system which enables the  
54  
55 screening of 120000 -240000 compounds per day has been created by the use of high  
56  
57  
58  
59  
60

1  
2  
3  
4 density 1536 wells plates. Higher capacity plates of 3456 wells are also available but  
5  
6 used rarely due to the limitation determined by the low liquid volume delivery  
7  
8 (Bajorath, 2002; Smith, 2002; Williams, 2004). Compounds found to be active in HTS  
9  
10 are called “hits”, but the distance from a hit to a lead is very great. This requires more  
11  
12 extensive screening for other activities, including toxicological evaluations. In this  
13  
14 process similar compounds will be produced to optimize the activity profile.  
15  
16  
17

18  
19 However, compounds obtained from the screening are not always (physico)  
20  
21 chemically and pharmacologically suitable for further development as a lead. While the  
22  
23 *in vitro* profile is maybe very promising, ADMET (absorption, distribution, metabolism,  
24  
25 elimination, toxicity) properties are often poor (Bleicher *et al.*, 2003). In most cases the  
26  
27 compound will require modifications to become a lead and drug candidate. Thus, the  
28  
29 crucial bottleneck of lead development is not only hit identification, but the potential of  
30  
31 the candidate to become a new drug (Bleicher *et al.*, 2003).  
32  
33  
34

35  
36 Therefore it is necessary not to focus only on quantity but also consider quality of  
37  
38 both library and assays. One suggestion made on this matter, was to introduce simple  
39  
40 ADMET parameters at the early stage of screening and the more sophisticated ones at a  
41  
42 later stage (Walters and Namchuk, 2003). The appropriate application of virtual  
43  
44 screening as a complement to a screening campaign to enrich the hit rate and reduce the  
45  
46 number of false positive has been suggested as well (Bajorath, 2002; Bleicher *et al.*,  
47  
48 2003; Entzeroth, 2003; Smith, 2002). Most virtual screening is guided by the structure  
49  
50 of the target protein and 3-dimensional descriptors of small molecular structures that fit  
51  
52 into the protein active site. Further approaches such as the identification of active  
53  
54 compounds by clustering and partitioning techniques, compound filtering to eliminate  
55  
56 undesired compounds (e.g. toxicant and groups with poor aqueous solubility and  
57  
58  
59  
60

1  
2  
3  
4 passive absorption) can also be applied (Bajorath, 2002). Understanding the mechanism  
5  
6 of action of a biological target is another approach. Several techniques have been  
7  
8 reported, such as immobilization and stabilization of the target protein upon its ligand  
9  
10 binding, as well as the use of nuclear magnetic resonance (NMR) spectroscopy coupled  
11  
12 to other analytical methods to learn about the drug-protein interaction (Entzeroth, 2003).  
13  
14 For the targets without sufficient bio-structural information, such as in the area of G-  
15  
16 protein coupled receptors (GPCRs), a ligand motif-based library design approach can be  
17  
18 applied (Bleicher *et al.*, 2003)  
19  
20  
21  
22

23  
24 The aforementioned efforts, however, did not successfully fulfill the high expectation  
25  
26 of increasing productivity. The tremendous amount of money spent was not  
27  
28 counterbalanced with the number of new approved drugs (Smith, 2002). However, all  
29  
30 these methods will result in all companies having the same basic information on targets  
31  
32 and probably ending up with the same compound, reducing the chance for real  
33  
34 innovation with this approach will be limited. This has encouraged reconsidering NP as  
35  
36 a drug source.  
37  
38  
39  
40  
41

### 42 **3. Efforts to combine natural products with new technologies**

43  
44 Many studies reported possible approaches which could potentially help to overcome  
45  
46 the above mentioned problems in the NP drug discovery process. Solutions to address  
47  
48 the incompatibility of NP extracts with HTS technology were proposed, including the  
49  
50 generation of a high quality NP compound library, the improvement of fractionation  
51  
52 methods, and of dereplication and identification steps.  
53  
54  
55  
56  
57  
58

#### 59 **3.1. Constructing a high quality natural product library**

60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
Several strategies to construct better NP libraries and to improve methods to screen these libraries have been reviewed. The advantages and limitations of a crude extract library, pre-fractionated library, or pure compound library have been described (Koehn, 2008). Crude extract libraries are simple to make, the obstacles appearing in the identification and dereplication steps. A pre-fractionated library can be effective to overcome these hurdles since the less complex fractions require fewer purification steps and the interference of unwanted compounds is reduced, while more investment is required both for fraction preparation and for the HTS assay since several samples are obtained from a single extract. The pure compound library is preferred since it is able to move forward directly after finding a hit.

28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
The strategy employed to build a fraction library as a compound source for HTS from Chinese herbal formulations reported to treat atherosclerosis has been described. Medium pressure liquid chromatography with a gradient solvent system followed by reverse phase semi-preparative HPLC was used as a filtering method to remove unfavorable compounds lacking desirable ADME properties and to prepare semi-crude extracts with different polarity. The fractions were characterized and identified by LC-DAD-ESI-MS. A component detection algorithm (CODA) was used to reduce noise and background of the chromatograms which then were presented as a histogram describing compound molecular distribution. Small molecules (less than 500 Da) were characterized by retention time, molecular weight, UV absorbance, and crude drug origin. In this way, 36 major compounds of each fraction were identified and after cell based multiples bioassay screening, validation and dereplication steps, 9 hits were found (Liu *et al.*, 2008).

### 3.2. Improving fractionation methods

Another study focused on sample fractionation using an automatic multiple solid phase extraction (SPE) method (Schmid *et al.*, 1999). Several different solid phases were used thus reducing the complexity of an extract from a microbial fermentation broth, leading to a more reliable and reproducible screening result.

### 3.3. Improving dereplication and identification steps

A fast dereplication and identification of novel active compounds can be achieved by direct coupling of the bioassay to liquid chromatography (Bajorath, 2002; van Elswijk and Irth, 2002). An example is the post-column acetylcholine esterase (AChE) inhibitor assay. After centrifugal partition chromatography (CPC) pre-fractionation and bioactivity testing, the most active *Narcissus* 'Sir Winston Churchill' fraction was injected into an online biochemical detection system coupling HPLC--UV-MS for AChE inhibitor activity. This system enabled a fast dereplication, allowing the identification of known compounds. It proved very useful for the detection of novel active compounds in the presence of known active compounds. For example, beside the well known galantamine, another active compound, unguiminorine, was identified (Ingkaninan *et al.*, 2000). A similar technique was applied to screen for natural phosphodiesterase and estrogen receptor ligands (Schenk *et al.*, 2003; Schobel *et al.*, 2001).

Sometimes, the information of molecular weight and UV spectra provided by combination of LC-DAD-MS is not sufficient for the identification of active compounds. The use of electrospray ionization Fourier Transform Ion Cyclotron Resonance MS (ESI-FTICR-MS) as a rapid and informative approach for direct affinity

1  
2  
3  
4 screening of bovine carbonic anhydrase II ( bCAII) ligands from extracts was reported  
5  
6 (Poulsen *et al.*, 2006). The bCAII-ligand affinity is based on the non-covalent binding  
7  
8 between the enzyme and the ligand. The electrospray ionization technique is known as a  
9  
10 gentle ionization process that can transfer intact non-covalent complexes from the  
11  
12 solution phase into the gas phase while the high resolution of the FTICR-MS allows the  
13  
14 very accurate deduction of mass-to-charge ratio (m/z) of the protein-ligand complex  
15  
16 from which even the molecular weight of this ligand can be calculated. This method was  
17  
18 applied by separating small molecules from macromolecules by size exclusion  
19  
20 chromatography (SEC), after the protein was added to the extract. This protein-ligand  
21  
22 complex was then subjected to online SEC-ESI-FTICR-MS or directly infused into ESI-  
23  
24 FTICR-MS, allowing the rapid determination of molecular weight of the active  
25  
26 compound with an error below 0.08 Da. Bioactive compounds which were present in  
27  
28 the mixture could be detected down to the level of 0.02% dry weight. HPLC coupled to  
29  
30 an on-line postcolumn antioxidant detection system in combination with absorbance  
31  
32 spectrometry and quadrupole time-of-flight high-resolution MS (QTOF-MS) allowed  
33  
34 the identification of compounds responsible for the antioxidant capacity of raspberry  
35  
36 (*Rubus idaeus*) and their level during raspberry ripening stages and in different cultivars  
37  
38 and different genetic background (Beekwilder *et al.*, 2005).  
39  
40  
41  
42  
43  
44  
45  
46

47 In a recent report, NMR based screening of bioactive NP and structure elucidation of  
48  
49 the active compound in a single step was described (Politi *et al.*, 2005). This study  
50  
51 aimed at detecting ligands from plants that bind to a specific target protein, lectin. The  
52  
53 interaction between a ligand present in the plant extract was detected by comparing the  
54  
55 saturation transfer difference (STD) and NOESY spectra of target protein with and  
56  
57 without extract. The most important structural information was provided by two  
58  
59  
60



1  
2  
3  
4 dimensional NMR data such as TOCSY, HSQC, HMBC and selective 1D-COSY.  
5  
6 Several LC-NMR hyphenation techniques and their application to advance plant extract  
7  
8 profiling have been described (Exarchou *et al.*, 2005). The advantage and limitation of  
9  
10 continuous flow LC-NMR, stop flow LC-NMR, LC-NMR/MS, LC-SPE-NMR, and  
11  
12 application of cryogenic and microflow technology have also been reviewed. The  
13  
14 performance of these LC-NMR coupling technologies depends on many factors such as  
15  
16 the physical connection of LC-NMR, flow-through probe design, and factors affecting  
17  
18 NMR sensitivity (solvent compatibility, solvent suppression, the adequate sample  
19  
20 volume). Practical examples of the application of various hyphenated LC-NMR  
21  
22 techniques for rapid screening, dereplication, and structure determination of active  
23  
24 compounds from natural extract has been published (Wolfender *et al.*, 2006).  
25  
26  
27  
28  
29  
30  
31  
32

### 33 **3.4. Virtual screening**

34  
35 The application of virtual screening of known NP has been proposed (Rollinger *et al.*,  
36  
37 2008). NP are still an incredible rich source of new drugs and because of the better  
38  
39 understanding of protein-ligand interaction properties, NP virtual screening can  
40  
41 constitute an alternative way to accelerate the drug discovery process. Many advantages  
42  
43 have been described including higher capacity, reduction in the number of experiments  
44  
45 for testing and isolating compounds, the ability to calculate the possible interaction of  
46  
47 compounds and targets, and the possibility of applying additional filters such as ADME  
48  
49 and drug-like properties to improve the hit quality. Strategies to combine virtual  
50  
51 screening and classical methods in NP exploration were proposed; (1) Direct screening  
52  
53 of a 3D database of NP using available target information and pharmacophore model of  
54  
55 specific activity, followed by the isolation of hits once the results are confirmed by  
56  
57  
58  
59  
60

1  
2  
3  
4 experimental data. (2) Extracts were roughly screened experimentally to select the  
5  
6 active ones, or the specific activity was explored through an ethnopharmacological  
7  
8 approach, the result of which was then virtually screened using a pharmacophore model  
9  
10 of the target, followed by isolation.  
11  
12

#### 13 14 15 16 **4. Metabolomics** 17

18  
19 Metabolomics is one of the new 'omics' fields, joining genomics, transcriptomics, and  
20  
21 proteomics (Rochfort, 2005). The general aim of metabolomics is the qualitative and  
22  
23 quantitative analysis of all metabolites (the metabolome), present in an organism, at a  
24  
25 specific time and under specific influence (Colquhoun, 2007; Hall, 2006; Mendes *et al.*,  
26  
27 2005; Ulrich-Merzenich *et al.*, 2007; Verpoorte *et al.*, 2007). A metabolome is the entire  
28  
29 set of small molecules present in a cell, tissue, or organism under certain conditions and  
30  
31 at certain point in time (Ulrich-Merzenich *et al.*, 2007).  
32  
33

34  
35 Metabolome components can be viewed as gene expression end products, thus its  
36  
37 quantitative and qualitative measurement may provide a broad insight of the  
38  
39 biochemical status of an organism and can be used to monitor gene function (Sumner *et*  
40  
41 *al.*, 2003). The use of transcriptomics and proteomics profiles to predict gene function  
42  
43 has been criticized because of limitations such as the fact that the changes in the  
44  
45 transcriptome or proteome do not always correlate to biochemical phenotypes or that the  
46  
47 translated protein may or may not be enzymatically active (Nobeli and Thornton, 2006;  
48  
49 Rochfort, 2005; Sumner *et al.*, 2003). The lack of a data base source may result in  
50  
51 limited information since the identification of mRNA and proteins is based on sequence  
52  
53 similarity or database matching. Therefore, among the 'omics' technologies,  
54  
55 metabolomics can be viewed as the most functional approach (Sumner *et al.*, 2003). The  
56  
57  
58  
59  
60

1  
2  
3  
4 combined information obtained from genomics, proteomics and metabolomics should  
5  
6  
7 be able to provide an integrated understanding of a cell or organism (Sumner *et al.*,  
8  
9 2003; Ulrich-Merzenich *et al.*, 2007).

#### 14 **4.1. Detection method**

16 Compared to proteins and RNA, secondary metabolites have a much more diverse  
17  
18 atomic arrangement which results in broad variations in chemical and physical  
19  
20 properties, such as molecular weight, polarity, solubility, and volatility (Dunn and Ellis,  
21  
22 2005), and can be present in a large range of concentrations (Moco *et al.*, 2006). This  
23  
24 generates technological difficulties in the simultaneous profiling of all metabolites  
25  
26 (Sumner *et al.*, 2003). There are thus considerable challenges in metabolomics  
27  
28 experiments such as sampling, sample preparation, instrumental analysis, and then in  
29  
30 the interpretation of the tremendous amount and diversity of data arising from the  
31  
32 metabolomics analysis (Bhalla *et al.*, 2005; Rochfort, 2005), the construction of  
33  
34 consolidated metabolite libraries and the development of metabolite-specific data  
35  
36 management systems (Moco *et al.*, 2006). Therefore, metabolomics based work still  
37  
38 depends on further improvement of sampling, analyzing and computational technologies  
39  
40 (Bhalla *et al.*, 2005; Rochfort, 2005).

47 The identification and quantification techniques applied to metabolomic based work  
48  
49 have been comprehensively reviewed (Bhalla *et al.*, 2005; Fiehn, 2002; Kell, 2004;  
50  
51 Sumner *et al.*, 2003; van der Greef and Smilde, 2005; Verpoorte *et al.*, 2008). Many  
52  
53 analytical methods with better reliability, robustness, selectivity, and resolution, have  
54  
55 been developed in which instead of targeting on a few specific compounds, the aim is to  
56  
57  
58  
59  
60

1  
2  
3  
4 identify and quantitate multiple targets in order to obtain an overview of compounds  
5  
6 classes (Fiehn, 2002).  
7

8  
9 The progressive development of ionization and mass detection methods, and the  
10 ability to detect a wide range of molecular weight and polarities have become a strong  
11 point for the use of MS and tandem MS techniques in metabolomics (Verpoorte *et al.*,  
12 2008). GC coupled to MS (GC-MS) has been used to identify and elucidate the  
13 structures of 326 different compounds from *Arabidopsis thaliana* leaf (Fiehn *et al.*,  
14 2000) and to identify and classify 322 tomato volatiles from 94 genotypes according to  
15 their biochemical precursor similarity (Tikunov *et al.*, 2005)  
16  
17  
18  
19  
20  
21  
22  
23  
24

25  
26 Liquid chromatography (LC) is an alternative tool particularly for non-volatile plant  
27 metabolites though it requires more technical support than GC (Dunn and Ellis, 2005;  
28 Nobeli and Thornton, 2006). An open access database dedicated to LC-MS based  
29 metabolomics of tomato fruit (*Solanum lycopersicum*) has been presented. A C18 base  
30 RP LC-photodiode array detection (PDA)-electrospray ionization (ESI)-quadrupole  
31 time-of-flight (QTOF)-MS method was developed for the detection and putative  
32 identification of predominantly semipolar secondary metabolites. Some previously  
33 unreported novel compounds from tomato fruit were identified by this means (Moco *et*  
34 *al.*, 2006).  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

47 The use of laser desorption/ionization mass spectrometry on porous silica (DIOS)  
48 and cluster analysis to discriminate the metabolic footprints of mutant yeast and wild  
49 type yeast has been reported (Vaidyanathan *et al.*, 2005). Compared to matrix- assisted  
50 laser desorption/ionization (MALDI) mass spectrometry, DIOS-MS is much simpler,  
51 allowing the analysis of small molecules in absence of matrix ions with little or no  
52 fragmentation and better tolerability to moderate amounts of contaminants in the  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4 biological sample. Both methods are also especially effective for analysis of proteins  
5  
6 and their derivatives (Thomas *et al.*, 2001).  
7  
8

9  
10 Although MS sensitivity is probably the best among analytical methods, there are  
11 some limitations for its application in metabolomics. For example, the different  
12 sensitivity of the detection of molecules and its reproducibility are important issues.  
13  
14 Because of different detector response, calibration curves are needed for each individual  
15  
16 compound for absolute quantification, this being totally unrealistic considering the  
17  
18 hundreds of signals that appear in a metabolomic profile. In this case, NMR  
19  
20 spectrometry is the method of choice, in absence of which only a relative  
21  
22 quantification of the individual compounds in different experiments can be made  
23  
24  
25  
26  
27  
28 (Verpoorte *et al.*, 2008)  
29

30  
31 The application of NMR for metabolic profiling of plant tissue extracts has been  
32  
33 reviewed by several authors (Colquhoun, 2007; Verpoorte *et al.*, 2007; Ward *et al.*,  
34  
35 2007). Reproducibility is the most important criteria for developing a metabolomics  
36  
37 technology platform, and NMR is the most suited choice even though its sensitivity is  
38  
39 not as high as MS (1  $\mu\text{M}$  – 1 mM in NMR tube) (Colquhoun, 2007; Verpoorte *et al.*,  
40  
41 2007). The simple and fast sample preparation, short measurement time, the availability  
42  
43 of less automatization and advanced data analysis methods, plus the possibility to  
44  
45 elucidate structures of known or unknown compounds in a complex mixture using  
46  
47 advanced two-dimensional (2D) NMR methods, are further advantages (Colquhoun,  
48  
49 2007; Verpoorte *et al.*, 2007).  
50  
51  
52  
53

## 54 55 56 **4.2. Data processing** 57 58 59 60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Data resulting from metabolomics based work are typically high dimensional data which require multivariate data analysis (MVDA) methods for interpretation. Most metabolomics data analysis methods are based on the classification of samples into different groups (e.g. treatment, genotype), both by supervised (e.g. discriminant function analysis, artificial neural networks) or unsupervised type data analysis method e.g. principal component analysis (PCA), hierarchical cluster analysis (HCA) (Nobeli and Thornton, 2006). It is also possible to use MVDA to make a regression modeling between two blocks of data, usually denoted as X and Y. In metabolomics based NP studies, X may represent signals from different metabolites present in plant extracts which are sampled at regular time intervals, and Y the responses (e.g. quality of product, bioactivity, yield). The model then can be used to predict Y from X which is obtained from new observations. The most common MVDA method for this modeling is partial least square (PLS) method (Wold *et al.*, 2001a). Some extended versions of PLS to improve the predictability power were developed recently, such as orthogonal signal correction (OSC), orthogonal-PLS (OPLS), and orthogonal-orthogonal-PLS (O2PLS) (Gabrielsson *et al.*, 2006; Wold *et al.*, 2001b)

## 5. Various applications of metabolomics in bioactivity screening of natural product

### 5.1. Metabolomics to improve identification and dereplication steps

The reductionist approach as described above is still widely applied in many NP drug discovery projects, i.e. in bioassay guided isolation work. In this approach, the rapid dereplication of known actives and identification of novel actives is essential. Addressing this issue, some papers have described the way in which a metabolomics approach consisting in the use of various analytical methods followed by appropriate

1  
2  
3  
4 MVDA can be used to shorten the reductionist bioassay driven isolation route,  
5  
6 especially in the identification and dereplication step.  
7  
8

9 In one study, the hierarchical cluster analysis of the chromatographic profiles of 500  
10 lipidic extracts of bacteria, resulted in the identification of 2 new  
11 oligoprenylsquaresquiterpenes and 2 new N-(2-arylethyl) amides, though the bioactivity  
12 profile was not reported here. HCA was chosen instead of PCA since the number of  
13 principal components representing more than 50% of the variance was much too high  
14 for PCA. PCA is more suitable for a large data set of similar compounds, which was not  
15 the case in this project. The use of this clustering method was able to advance the  
16 dereplication step and to avoid redundant analyses for selection of strains containing  
17 new NP (Böröczky *et al.*, 2006).  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29

30 The most time consuming step in metabolomics based work is the identification of  
31 metabolites in a prefractionated mixture. The availability of a compound spectral library  
32 is crucial for rapid identification. A free web-based tool containing data related to  
33 primary and secondary metabolites from various species based on NMR or MS data, the  
34 Madison Metabolomics Consortium Database ((MMCD) was developed (Cui *et al.*,  
35 2008). The sensitivity of MS and the feasibility of NMR spectroscopy, including the  
36 application of two dimensional (2D) NMR followed by MVDA application, make these  
37 techniques particularly well suited for the characterization of crude unfractionated NP.  
38  
39 The analysis based on double quantum filtered correlation spectroscopy (DQF-COSY)  
40 was introduced to simplify the analysis of 2D NMR spectra and has demonstrated its  
41 utility to detect previously unreported terpenoid indole alkaloids from *Tolypocladium*  
42 *cylindrosporum* grown in different media. Graphical analysis of the obtained spectra  
43 through multiplicative stacking and graphical manipulation of the bitmap spectra  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4 allowed signals only present in a specific media to be distinguished. Further structure  
5  
6 elucidation was then performed by heteronuclear single quantum correlation (HSQC)  
7  
8 and heteronuclear multiple bond correlation (HMBC) spectra analysis. This method is  
9  
10 considerably more efficient than spectra subtraction since results in a clear color  
11  
12 alteration and partially obsolescence of a common signals present in two spectra even if  
13  
14 those are present in a distinct different concentration (Schroeder *et al.*, 2007). The big  
15  
16 size of data limits the application of this method. Three different clustering methods, i.e.  
17  
18 tree clustering analysis (TCA), K-means clustering and multidimensional scaling, were  
19  
20 used to cluster together partially purified plant and marine organism extracts which  
21  
22 contain the same bioactive compound or groups of compounds. These clustering  
23  
24 methods were chosen to overcome limitations of PCA to handle the data with a large  
25  
26 number of clusters or subsets within the data (Pierens *et al.*, 2005). Only a few of the  
27  
28 secondary metabolites from *Myxococcus xanthus* have been investigated, in particular  
29  
30 some antibiotics such as myxovirescin, saframycin, and althiomycin. However, only  
31  
32 one metabolite unique to *M. xanthus* was known, that is DKxanthanes. Intraspecific  
33  
34 diversity in the metabolite profiles of 98 *Myxococcus xanthus* strains originated from  
35  
36 78 locations worldwide was rapidly explored by means of fast chromatographic  
37  
38 separation (UPLC)-coupled ESI-TOF-MS and PCA. As a result, 37 candidates for novel  
39  
40 compounds were identified but their bioactivities were not reported (Krug *et al.*, 2008).  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50

## 51 **5.2. Metabolomics for quality control of phytomedicines**

52  
53 At present, the most popular application of metabolomics in NP research is for the  
54  
55 quality control of phytomedicines. Different metabolite profiles may occur due to  
56  
57 variations between species or variety, adulteration, environmental changes during  
58  
59  
60



1  
2  
3  
4 growth and harvesting, post harvesting treatment, extraction and method of preparation.  
5  
6 These factors may significantly alter the efficacy of phytomedicines prepared with  
7 these herbs (Wang *et al.*, 2005). The application of  $^1\text{H}$  NMR combined with  
8 multivariate data analysis to detect variations between different herbal preparations has  
9 been reported. The most common methods that have been applied for these ends are  
10 unsupervised principal component analysis (PCA) and supervised partial least square  
11 analysis/ partial least square analysis with discriminant analysis (PLS/PLS-DA). These  
12 methods were applied to differentiate various chamomile (*Matricaria recutita* L.)  
13 preparations made with plants grown in different locations, percentage of the desired  
14 plant part, and extraction method (Wang *et al.*, 2004). Discrimination between different  
15 *Cannabis sativa* cultivars based on the content of  $\Delta^9$ -tetrahydrocannabinolic acid  
16 (THCA) and cannabidiolic acid (CBDA) was achieved (Choi *et al.*, 2004), as well as  
17 discrimination between 11 *Ilex* species including 3 varieties (Choi *et al.*, 2005).  
18 Similarly, discrimination between *Ephedra* species (Kim *et al.*, 2005), and commercial  
19 feverfew (*Tanacetum parthenium*) preparations from different suppliers (Bailey *et al.*,  
20 2002) was obtained. The discrimination of commercial preparations of St. John's Wort  
21 (*Hypericum perforatum*) from different batches of the same supplier, including variation  
22 of flavonoid content which have been previously related to the antidepressant activity  
23 was also reported (Rasmussen *et al.*, 2006). An example of application of GC and GC-  
24 MS combined with multivariate data analysis to correlate artemisinin concentration with  
25 the growth stage of *Artemisia annua* L. and to distinguish between transgenic and wild  
26 type has been reported (Ma *et al.*, 2008).  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55

56 Metabolomics can also be used to detect adulterations of herbal preparations with  
57 similar species but with low levels of the desired active compound(s).  $^1\text{H}$  NMR and  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

PCA was used to verify that the anti-malaria capsules produced by one specific company did not contain *Artemisia annua* as labeled but *Artemisia afra*, which has never been reported to contain the active compound artemisinin or any of its derivatives (van der Kooy *et al.*, 2008). Instead of  $^1\text{H}$  NMR, 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 erythrosa 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).

### 5.3. Metabolomics to link chemical profile and bioactivity pattern of phytomedicines

As previously described, a reductionist approach focusing on a single active compound is necessary in some cases, but in case of St. John’s Wort extract, where no single compound or even group of compounds have been found to be responsible for its activity, the possibility of synergism and pro-drugs is evident, and a holistic *in vivo* approach would be more valuable to guarantee its therapeutic efficacy (Ulrich-Merzenich *et al.*, 2007; Verpoorte *et al.*, 2005). Animal experiments and clinical trials are two classical ways to conduct *in vivo* experiments, but the time required and costs involved make this option practically non viable. Furthermore, if these tests are performed with extracts of uncertain composition, results will invariably be unreliable or inconsistent. Recently new approaches are evolving which might help to get a better view of the mode of action of herbal or natural products and lead to the possibility of

1  
2  
3  
4 obtaining proof and guarantee of their pharmacological activity over different batches.  
5  
6 These approaches are based on considering the NP holistically, i.e. on the bulk of  
7  
8 chemical constituents rather than focusing on certain individual components or groups  
9  
10 of compounds. On the other hand, the changes caused by a sample on the transcriptome,  
11  
12 proteome and/or metabolome patterns can be compared with those observed after  
13  
14 treatment with known drugs rather than looking only at certain specific changes.  
15  
16 Genomics, proteomics, and metabolomics platforms will be crucial since they allow the  
17  
18 identification of detectable protein/genes or compounds which might relate to the  
19  
20 biological activity of complex extract (Ulrich-Merzenich *et al.*, 2007). This may shift  
21  
22 the paradigm in the development and application of complex plant/phytocompound  
23  
24 mixtures in modern medicine (Shyur and Yang, 2008). Moreover, by measuring the  
25  
26 effect of a complex extract administered to a living organism, either *in vitro* or *in vivo*, it  
27  
28 is possible to identify the compound or combination of compounds correlated with  
29  
30 activity. Besides, new or unsuspected mode(s) of action and targets might also be  
31  
32 revealed in this approach (Verpoorte *et al.*, 2005).  
33  
34  
35  
36  
37  
38  
39

40 Several papers discussed a metabonomics approach to study various aspects of  
41  
42 natural products extracts such as bioavailability, toxicity, and proof of efficacy.  
43  
44 Metabonomics is defined as ‘the quantitative measurement of the dynamic  
45  
46 multiparametric metabolic response of living systems to pathophysiological stimuli or  
47  
48 genetic modification’ (Nicholson *et al.*, 2002). This term is in general equivalent to the  
49  
50 NMR spectroscopic analysis of biofluids to identify changes in the concentrations and  
51  
52 patterns of endogenous metabolites induced by exogenic compounds (Nobeli and  
53  
54 Thornton, 2006) but mostly used in medicinal or toxicological studies although  
55  
56 sometimes the term metabolomics is also used.  
57  
58  
59  
60

1  
2  
3  
4 A summary on the implementation of metabolomics, particularly NMR based  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

metabolomics coupled to multi-variate data analysis tools to detect links between the chemical profile of a herbal preparation and its bioactivity is provided in Figure 1. A quick review of the literature reflects the enormous amount of studies that are carried out nowadays using this technology, some representative examples of which are summarized below.

### 5.3.1. Identification of active compounds and quantitative prediction of bioactivity

Chemometric and  $^1\text{H}$  NMR analyses were used to discriminate *Artemisia annua* L samples obtained from different sources. PCA assignment showed that artemisinin is the main discriminant factor for clustering. PLS analysis was applied to predict the quantitative anti-plasmodial activity ( $\text{IC}_{50}$ ) and toxicity ( $\text{ToxEd}_{50}$ ) value of different *A. annua* extracts. The predictive values were in close agreement with the actual experimental values (Bailey *et al.*, 2004). The same approach was used to demonstrate the correlation between the spectral pattern and opioid receptor binding properties of St. John's Wort (*Hypericum perforatum*) extracts prepared with different solvents (Roos *et al.*, 2004). 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 to *Z. officinale*, but there was lack of correlation between the content of gingerol, a compound specific to *Z. officinale*, and the observed total anti-inflammatory activity. Therefore, it was concluded that other compounds, unidentified to date which contributed to this activity had to be detected and quantified in order to guarantee the bioactivity of a particular *Zingiber* sample (Jiang *et al.*, 2006). The application of  $^1\text{H}$  NMR and PCA analysis allowed the differentiation of six samples of

1  
2  
3  
4 *Galphimia glauca*, a Mexican herbal used to treat central nervous disorders, collected  
5  
6 from six different areas. Galphimines and 1,3,4,5-tetra-*O*-galloylquinic acid contents  
7  
8 were found to be the main discriminating factors for PCA separation. Further analysis  
9  
10 with PLS-DA confirmed that the level of both compounds was correlated with sedative  
11  
12 and anxiolytic effects in mice (Figure 2) (Cardoso-Taketa *et al.*, 2008). This technique  
13  
14 was also applied successfully to distinguish thrips-resistant and –susceptible F2 hybrids  
15  
16 of *Senecio jacobaea* and *Senecio aquaticus*. Pyrrolizidine alkaloids, kaempferol and  
17  
18 jacaranone were found to be higher in resistant plants. This is of interest since the last  
19  
20 two compounds have positive effects on human health as anti-cancer agents (Leiss *et*  
21  
22 *al.*, 2009).

### 23 24 25 26 27 28 **5.3.2. Proof of efficacy and mode of action identification of herbal medicines**

29  
30 The *in vitro* cellular effects of *Cannabis* from different cultivars and preparations  
31  
32 were analyzed using a proteomic and transcriptomic platform (Wang *et al.*, 2005). In  
33  
34 another case, the *in vivo* effect of *Ginkgo* on rats was studied by applying a <sup>1</sup>H-NMR  
35  
36 and supervised MVDA method to measure metabolites in their urine (Wang *et al.*,  
37  
38 2005). The effect of three traditional Chinese medicines on insulin resistance APOE\*3-  
39  
40 Leiden transgenic mice which are highly responsive in their plasmatic lipids to dietary  
41  
42 treatments was also studied. LC-MS measurement combined with MVDA was applied  
43  
44 to three TCM preparations to discover which was most effective TCM but had the  
45  
46 lowest toxicity (Wang *et al.*, 2005).

47  
48  
49  
50  
51 The use of metabolomic profiling to study the mode of action of active natural  
52  
53 products recently emerged as well, predominantly in antimicrobial activity analysis. As  
54  
55 an example, HPLC/DAD/ESI-MS application followed by PCA confirmed that  
56  
57 dihydrocurbitacin F-25-*O*-acetate was the main component having an antimicrobial  
58  
59  
60

1  
2  
3  
4 activity in *Hemsleya pengxianensis*. The possible antibacterial mode of action on  
5  
6  
7 *Staphylococcus aureus* CCTCC AB9105 was predicted by comparing the metabolite  
8  
9  
10 profile of *S. aureus* treated by the plant extract and commercial antibacterials with  
11  
12 known mode of action (Biao-Yi *et al.*, 2008).

13  
14 The use of metabonomics to prove the efficacy of herbal medicines was reported in  
15  
16 some papers. The evaluation of the anti-ageing effects of the total flavone content of  
17  
18 Herba Epimedii (TFE), a famous Chinese herbal medicine used to treat some age-  
19  
20 related diseases was conducted by <sup>1</sup>H NMR analysis and PCA assignment. Ten  
21  
22 metabolites were identified as ageing markers in the rat urine profile: creatinine, lactate,  
23  
24 alanine, acetate, acetone, succinate, allantoin, methylamine, dimethylamine and  
25  
26 trimethylamine-N-oxide. It showed that after TFE intervention, the 24-month-old rats  
27  
28 urine profile was shifted towards that of 18-months-old rats (Wu *et al.*, 2008). In  
29  
30 another case, the study of *Epimedium brevicornum* Maxim, was undertaken. This herbal  
31  
32 medicine, one of the most popular in China is used for toning the kidney and  
33  
34 strengthening bones, but the kind of constituents responsible for its pharmacological  
35  
36 effect, and the mechanism through which it affects the entire body metabolism was  
37  
38 unknown. The plasma metabolites of rats from three different treatments (pre-  
39  
40 hydrocortisone intervention, post-hydrocortisone intervention, and *Epimedium*  
41  
42 *brevicornum* Maxim. treated) were analyzed using UPLC-MS followed by PCA. The  
43  
44 distinct increase of the level of three metabolites was found after hydrocortisone  
45  
46 intervention, one of them identified as ethylindole-3-crylic acetate. The result showed  
47  
48 that after *Epimedium brevicornum* Maxim treatment, the plasma metabolite profile was  
49  
50 restored closer to that of the pre-hydrocortisone intervention. Of the four compounds of  
51  
52 this herbal preparation detected in plasma (epimedin C, icariin, icariside II, and 2''-O-  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4 rhamnosoyl icaride), epimedin C and icariin were also found in urine, and it was  
5  
6 concluded that those might be the active ones (Famei Li *et al.*, 2007).  
7  
8

9  
10 The efficacy and the mechanism of action of the traditional Chinese medicine  
11 preparation known as Xindi soft capsules (consisting of sea buckthorn flavonoids and  
12 sea buckthorn berry oil) used to treat blood stasis, was studied by UPLC combined with  
13  
14 quadrupole time-of-flight tandem mass spectrometry (UPLC-QTOF-MS) followed by  
15  
16 PCA and PLSDA analysis. Blood stasis includes hematological disorders such as  
17  
18 haemorrhage, nasal congestion, thrombosis, and local ischemia (microclots).  
19  
20 The comparison of the urinary metabolite pattern obtained for 5 groups of rats (healthy  
21 control group, acute blood stasis model group, low dose group of Xindi soft capsule,  
22 middle dose group of Xindi soft capsule, high dose group of Xindi soft capsule),  
23 showed that the treated groups were located between the acute blood stasis model group  
24 and the healthy control group. Confirmation by hemorheological analysis proved that  
25 within the treated group, especially in the case of the highest dose group, the urine  
26 metabolite pattern tends toward that of the control group. Some potential biomarkers  
27 were identified as well, such as cholic acid, phenylalanine and kynurenic acid (Zhao *et*  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

### 5.3.3. Bioavailability and fate of natural compounds assessment

The metabonomics approach was successfully applied to study the bioavailability and fate of natural compounds. A traditional Chinese medicine formulation, Dangguibuxue decoction (ODD) consisting of huangqi (*Radix astragali*)-danggui (*Radix angelicae sinensis*), is widely used as a tonic. ODD is a very complex mixture containing hundreds of components. More than 70 compounds have currently been isolated and identified but only a few constituents were reported to be responsible for its



1  
2  
3  
4 activity. Metabolic fingerprinting by LC/DAD-MS was suitable for analysis of the  
5  
6 bioavailability of the metabolites of the ODD components *in vivo* by comparing  
7  
8 LC/DAD-MS fingerprint of the ODD formula with the metabolite fingerprint of plasma  
9  
10 of rabbits treated with ODD. Identification of detected compounds was done by MS and  
11  
12 UV spectra comparison with references from literature. Some flavonoids, phtalides, and  
13  
14 triterpene saponins, which were known as ODD active compounds were detected,  
15  
16 showing that they are well absorbed into the body (Wang *et al.*, 2007).  
17  
18  
19

#### 20 21 **5.3.4. Safety and toxicity of herbal medicines assessment**

22  
23 There are also many reports on the application of the metabonomics approach to  
24  
25 solve the two other major issues related to the quality control of herbal medicines -  
26  
27 safety and toxicity. The mechanism of the toxicity of aristolochic acid was studied by  
28  
29 analyzing the urine of rats submitted to intra-peritoneal injection of aristolochic acid and  
30  
31 several toxins with known modes of action. The PCA, confirmed by biochemical  
32  
33 parameter analysis, allowed researchers to conclude that aristolochic acid induced the  
34  
35 damage of renal proximal tubules and medulla, accompanied with a slight hepatic lesion  
36  
37 (Zhang *et al.*, 2006). Similarly, the toxicity of Hei-Shun-Pian (the processed lateral root  
38  
39 of *Aconitum carmichaelii* Debx., Ranunculaceae) was evaluated using the <sup>1</sup>H NMR  
40  
41 spectral metabolic profile of urine and plasma samples of rats after 5 days oral  
42  
43 treatment with the decoction of Hei-Shun-Pian at different doses. Hei-Shun-Pian is a  
44  
45 traditional Chinese medicine with analgesic, anti-pyretic, anti-rheumatoid arthritis and  
46  
47 anti-inflammatory effects. The C-19 diterpenoid alkaloid aconitine and a series of its  
48  
49 derivatives are suggested to be the active principles, but some of those are also very  
50  
51 toxic and have narrow therapeutic windows (LD50/ED50). Studying the PCA and PLS  
52  
53 analysis of <sup>1</sup>H NMR data, it was found that the primary differences between the plasma  
54  
55  
56  
57  
58  
59  
60



1  
2  
3  
4 of treated and the control groups were the increase in the concentration of lipids and the  
5  
6 decrease in the concentration of phosphatidylcholine (PtCho) and *O*- and *N*-acetyl  
7  
8 glycoproteins, while taurine, acetate, creatinine, 2-oxoglutarate, dimethylamine and  
9  
10 hippurate levels in urine contributed to the distinction between high and medium dose  
11  
12 groups and control group (Li *et al.*, 2008a).  
13  
14  
15  
16  
17

### 18 **Conclusions and perspectives**

19  
20  
21  
22  
23 In response to the renewed interest in NP, several strategies have been designed to  
24  
25 improve NP lead finding. Strategies aimed at improving the compatibility of natural  
26  
27 products with high throughput screening requirements by improving the NP libraries  
28  
29 shorten the discovery route, using advanced analytical methods and combining them  
30  
31 with online bioassays, and selecting new approaches which can rapidly provide  
32  
33 evidence of efficacy.  
34  
35  
36

37 The reductionist approach is still widely used in NP drug discovery studies.  
38  
39 However, metabolomics coupled with MVDA can be applied in those cases to speed up  
40  
41 the limiting steps in dereplication and identification of active compounds. This must be  
42  
43 supported by the availability of a high quality open access NP library based on NMR or  
44  
45 MS data. Metabolomics is undoubtedly an important tool for the quality control of  
46  
47 herbal preparations. It captures the differences in the composition due to different  
48  
49 cultivars, growth areas, methods of preparation, and adulterations, all of which might  
50  
51 significantly affect the content of active principles in the preparation.  
52  
53  
54  
55

56 To explain the chemodiversity offered by Nature in a more intelligent and  
57  
58 innovative way than random studies using HTS can be employed. For this systems  
59  
60

1  
2  
3  
4 biology based approach, metabolomics is viewed as a crucial tool. Systems biology may  
5  
6 allow the connection between the chemical profile of herbal preparations and the  
7  
8 response given by a living system, either *in vitro* or *in vivo*. Such a holistic approach to  
9  
10 obtain proof of efficacy has emerged as the new paradigm to detect novel NP in drug  
11  
12 discovery. Significant and obvious advantages are that synergistic interaction can be  
13  
14 detected, as well as prodrugs. At the same time metabolomics will provide information  
15  
16 on the mode of action and include toxic side effects. The capability to quickly analyze a  
17  
18 broad range of metabolites makes metabolomics important in the growing field of  
19  
20 personalized medicine. It allows the metabolic profile of individual patients to be linked  
21  
22 to their response to specific drug therapy as a base to develop a concept on markers of  
23  
24 the efficacy, toxicity, and clearance of drugs, permitting a better prescription for an  
25  
26 individual patient based on their genetic makeup and phenotype status.  
27  
28  
29  
30  
31  
32

33 It is evident that metabolomics has shown to be a potent facilitator in the recovery of  
34  
35 NP as excellent sources for novel leads, and even further, as a means to discover novel  
36  
37 targets, as the strength of the “omics” based approaches lies greatly in the fact that  
38  
39 results can be stored for long term data mining.  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## REFERENCES

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- Bailey NJC, Sampson J, Hylands PJ *et al.* 2002. Multi-component metabolic classification of commercial feverfew preparations via High-Field  $^1\text{H}$ -NMR spectroscopy and chemometrics. *Planta Med* **68**:734-738.
- Bailey NJC, Wang Y, Sampson J *et al.* 2004. Prediction of anti-plasmodial activity of *Artemisia annua* extracts: application of  $^1\text{H}$  NMR spectroscopy and chemometrics. *J Pharm Biomed Anal* **35**:117-126.
- Bajorath J. 2002. Integration of virtual and high-throughput screening. *Nat Rev Drug Discov* **1**:882-894.
- Beekwilder J, Jonker H, Meesters P *et al.* 2005. Antioxidants in raspberry; On-line analysis links antioxidant activity to a diversity of individual metabolites. *Agric Food Chem* **53**:3313-3320.
- Bhalla R, Narasimhan K, Swarup S. 2005. Metabolomics and its role in understanding cellular responses in plants. *Plant Cell Rep* **24**:562-571.
- 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.
- Bindseil KU, Jakupovic J, Wolf D *et al.* 2001. Pure compound libraries; a new perspective for natural product based drug discovery. *Drug Discov Today* **6**:840-847.
- Bleicher KH, Bohm H-J, Muller K *et al.* 2003. Hit and lead generation: beyond high-throughput screening. *Nat Rev Drug Discov* **2**:369-378.

- 1  
2  
3  
4 Böröczky K, Laatsch H, Döbler IW *et al.* 2006. Cluster analysis as selection and  
5  
6 dereplication tool for the identification of new natural compounds from large  
7  
8 sample sets. *Chem Biodiversity* **3**:622-634.  
9  
10  
11 Butler MS. 2004. The role of natural product chemistry in drug discovery. *J Nat Prod*  
12  
13 **67**:2141-2153.  
14  
15  
16 Cardoso-Taketa AT, Pereda-Miranda R, Choi YH *et al.* 2008. Metabolic profiling of  
17  
18 the Mexican anxiolytic and sedative plant *Galphimia glauca* using Nuclear  
19  
20 Magnetic Resonance spectroscopy and Multivariate Data Analysis. *Planta Med*  
21  
22 **74**:1295-1301.  
23  
24  
25  
26 Choi YH, Kim HK, Hazekamp A *et al.* 2004. Metabolomic differentiation of  
27  
28 *Cannabis sativa* cultivars using <sup>1</sup>H NMR spectroscopy and Principal Component  
29  
30 Analysis. *J Nat Prod* **67**:953-957.  
31  
32  
33 Choi YH, Sertic S, Kim HK *et al.* 2005. Classification of *Ilex* species based on  
34  
35 metabolomic fingerprinting using Nuclear Magnetic Resonance and Multivariate  
36  
37 Data Analysis. *J Agric Food Chem* **53**:1237-1245.  
38  
39  
40 Colquhoun IJ. 2007. Use of NMR for metabolic profiling in plant systems. *J Pestic*  
41  
42 *Sci* **32**:200-212.  
43  
44  
45 Cui Q, Lewis IA, Hegeman AD *et al.* 2008. Metabolite identification via the  
46  
47 Madison Metabolomics Consortium Database. *Nat Biotech* **26**:162-164.  
48  
49  
50 Dickson M, Gagnon JP. 2004. Key factors in the rising cost of new drug discovery  
51  
52 and development. *Nat Rev Drug Discov* **3**:417-429.  
53  
54  
55 Dunn WB, Ellis DI. 2005. Metabolomics: Current analytical platforms and  
56  
57 methodologies. *Trends Anal Chem* **24**:285-294.  
58  
59  
60

- 1  
2  
3  
4 Entzeroth M. 2003. Emerging trends in high-throughput screening. *Curr Opin*  
5  
6 *Pharmacol* **3**:522-529.  
7  
8  
9 Exarchou V, Krucker M, van Beek TA *et al.* 2005. LC-NMR coupling technology:  
10  
11 recent advancements and applications in natural products analysis. *Magn Reson*  
12  
13 *Chem* **43**:681-687.  
14  
15  
16 Famei Li, Xiumei Lu, Huiping Liu *et al.* 2007. A pharmaco-metabonomic study on  
17  
18 the therapeutic basis and metabolic effects of *Epimedium brevicornum* Maxim on  
19  
20 hydrocortisone-induced rat using UPLC-MS. *Biomed Chromatogr* **21**:397-405.  
21  
22  
23 Feher M, Schmidt JM. 2003. Property distributions: Differences between drugs,  
24  
25 natural products, and molecules from combinatorial chemistry. *J Chem Inf*  
26  
27 *Comput Sci* **43**:218-227.  
28  
29  
30 Fiehn O. 2002. Metabolomics – the link between genotypes and phenotypes. *Plant*  
31  
32 *Mol Biol* **48**:155-171.  
33  
34  
35 Fiehn O, Kopka J, Trethewey RN *et al.* 2000. Identification of uncommon plant  
36  
37 metabolites based on calculation of elemental compositions using Gas  
38  
39 Chromatography and Quadrupole Mass Spectrometry. *Anal Chem* **72**:3573-3580.  
40  
41  
42 Gabrielsson J, Jonsson H, Airiau C *et al.* 2006. OPLS methodology for analysis of  
43  
44 pre-processing effects on spectroscopic data. *Chemom Intell Lab Syst* **84**:153-  
45  
46 158.  
47  
48  
49 Gómez-Hens A, Aguilar-Caballos MP. 2007. Modern analytical approaches to high-  
50  
51 throughput drug discovery. *Trends Anal Chem* **26**:171-182.  
52  
53  
54 Hall RD. 2006. Plant metabolomics: from holistic hope, to hype, to hot topic. *New*  
55  
56 *Phytol* **169**:453-468.  
57  
58  
59  
60

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- Henkel T, Brunne RM, Müller H *et al.* 1999. Statistical investigation into the structural complementarity of natural products and synthetic compounds. *Angew Chem Int Ed* **38**:643-647.
- Ingkaninan K, Hazekamp A, de Best CM *et al.* 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.
- Jahangir M, Kim HK, Choi YH *et al.* 2008. Metabolomic response of *Brassica rapa* submitted to pre-harvest bacterial contamination. *Food Chemistry* **107**:362-368.
- Jiang H, Xie Z, Koo HJ *et al.* 2006. Metabolic profiling and phylogenetic analysis of medicinal *Zingiber* species: Tools for authentication of ginger (*Zingiber officinale* Rosc.). *Phytochemistry* **67**:1673-1685.
- Kell DB. 2004. Metabolomics and systems biology: making sense of the soup. *Curr Opin Microbiol* **7**:296-307.
- Kim HK, Choi YH, Erkelens C *et al.* 2005. Metabolic fingerprinting of *Ephedra* species using <sup>1</sup>H-NMR spectroscopy and Principal Component Analysis. *Chemical & Pharmaceutical Bulletin* **53**:105-109.
- Koehn FE. 2008. High impact technologies for natural products screening. *Natural Compounds as Drugs Volume I*. p 175-210.
- Koehn FE, Carter GT. 2005. The evolving role of natural products in drug discovery. *Nat Rev Drug Discov* **4**:206-220.
- Krug D, Zurek G, Revermann O *et al.* 2008. Discovering the hidden secondary metabolome of *Myxococcus xanthus*: a study of intraspecific diversity. *Appl Environ Microbiol*. **74**:3058-3068.

- 1  
2  
3  
4 Landro JA, Taylor ICA, Stirtan WG *et al.* 2000. HTS in the new millennium: The  
5  
6 role of pharmacology and flexibility. *J Pharmacol Toxicol Methods* **44**:273-289.  
7  
8  
9 Leiss K, Choi Y, Abdel-Farid I *et al.* 2009. NMR metabolomics of thrips (  
10  
11 *Frankliniella occidentalis* ) resistance in *Senecio* hybrids. *J Chem Ecol* **35**:219-  
12  
13 229.  
14  
15  
16 Li L, Sun B, Zhang Q *et al.* 2008a. Metabonomic study on the toxicity of Hei-Shun-  
17  
18 Pian, the processed lateral root of *Aconitum carmichaelii* Debx. (Ranunculaceae).  
19  
20 *J Ethnopharmacol* **116**:561-568.  
21  
22  
23 Li P, Qi L-W, Liu EH *et al.* 2008b. Analysis of Chinese herbal medicines with  
24  
25 holistic approaches and integrated evaluation models. *Trends Anal Chem* **27**:66-  
26  
27 77.  
28  
29  
30 Liu L, Li Y-F, Cheng Y-Y. 2008. A method for the production and characterization  
31  
32 of fractionated libraries from Chinese herbal formulas. *J Chromatogr B* **862**:196-  
33  
34 204.  
35  
36  
37 Ma C, Wang H, Lu X *et al.* 2008. Metabolic fingerprinting investigation of  
38  
39 *Artemisia annua* L. in different stages of development by gas chromatography and  
40  
41 gas chromatography-mass spectrometry. *J Chromatogr A* **1186**:412-419.  
42  
43  
44 Macarron R. 2006. Critical review of the role of HTS in drug discovery. *Drug*  
45  
46 *Discov Today* **11**:277-279.  
47  
48  
49 McChesney JD, Venkataraman SK, Henri JT. 2007. Plant natural products: Back to  
50  
51 the future or into extinction? *Phytochemistry* **68**:2015-2022.  
52  
53  
54 Mendes P, Camacho D, de la Fuente A. 2005. Modelling and simulation for  
55  
56 metabolomics data analysis. *Biochem Soc Trans* **33**:1427-1429.  
57  
58  
59  
60

- 1  
2  
3  
4 Mishra KP, Ganju L, Sairam M *et al.* 2008. A review of high throughput technology  
5  
6 for the screening of natural products. *Biomed Pharmacother* **62**:94-98.  
7  
8  
9 Moco S, Bino RJ, Vorst O *et al.* 2006. A Liquid Chromatography-Mass  
10  
11 Spectrometry-Based Metabolome Database for Tomato. *Plant Physiol* **141**:1205-  
12  
13 1218.  
14  
15  
16 Myers PL. 1997. Will combinatorial chemistry deliver real medicines? *Curr Opin*  
17  
18 *Biotechnol* **8**:701-707.  
19  
20  
21 Newman DJ, Cragg GM. 2007. Natural products as sources of new drugs over the  
22  
23 last 25 years. *J Nat Prod* **70**:461-477.  
24  
25  
26 Newman DJ, Cragg GM, Snader KM. 2003. Natural products as sources of new drugs  
27  
28 over the period 1981-2002. *J Nat Prod* **66**:1022-1037.  
29  
30  
31 Nicholson JK, Connelly J, Lindon JC *et al.* 2002. Metabonomics: a platform for  
32  
33 studying drug toxicity and gene function. *Nat Rev Drug Discov* **1**:153-161.  
34  
35  
36 Nisbet LJ, Moore M. 1997. Will natural products remain an important source of drug  
37  
38 research for the future? *Curr Opin Biotechnol* **8**:708-712.  
39  
40  
41 Nobeli I, Thornton JM. 2006. A bioinformatician's view of the metabolome.  
42  
43 *Bioessays* **28**:534-545.  
44  
45  
46 Ortholand J-Y, Ganesan A. 2004. Natural products and combinatorial chemistry:  
47  
48 back to the future. *Curr Opin Chem Biol* **8**:271-280.  
49  
50  
51 Pierens GK, Palframan ME, Tranter CJ *et al.* 2005. A robust clustering approach for  
52  
53 NMR spectra of natural product extracts. *Magn Reson Chem* **43**:359-365.  
54  
55  
56 Pieters L, Vlietinck AJ. 2005. Bioguided isolation of pharmacologically active plant  
57  
58 components, still a valuable strategy for the finding of new lead compounds? *J*  
59  
60 *Ethnopharmacol* **100**:57-60.



- 1  
2  
3  
4 Politi M, Chávez MI, Cañada FJ *et al.* 2005. Screening by NMR: A new approach  
5  
6 for the study of bioactive natural products? The example of *Pleurotus ostreatus*  
7  
8 hot water extract. Eur J Org Chem **2005**:1392-1396.  
9  
10  
11 Poulsen S, Davis R, Keys T. 2006. Screening a natural product-based combinatorial  
12  
13 library using FTICR mass spectrometry. Bioorg Med Chem **14**:510-515.  
14  
15  
16 Rasmussen B, Cloarec O, Tang H *et al.* 2006. Multivariate analysis of integrated and  
17  
18 full-resolution <sup>1</sup>H-NMR spectral data from complex pharmaceutical preparations:  
19  
20 St. John's Wort. Planta Med **72**:556-563.  
21  
22  
23 Rochfort S. 2005. Metabolomics reviewed: A new omics platform technology for  
24  
25 systems biology and implications for natural products research. J Nat Prod  
26  
27 **68**:1813-1820.  
28  
29  
30 Rollinger JM, Stuppner H, Langer T. 2008. Virtual screening for the discovery of  
31  
32 bioactive natural products. In: Petersen F, Amstutz R, editors. Natural  
33  
34 Compounds as Drugs Volume I. Basel: Birkhäuser Verlag p211-249.  
35  
36  
37 Roos G, Röseler C, Büter KB *et al.* 2004. Classification and correlation of St. John's  
38  
39 Wort extracts by Nuclear Magnetic Resonance spectroscopy, Multivariate Data  
40  
41 Analysis and pharmacological activity. Planta Med **70**:771-777.  
42  
43  
44 Schenk T, Breel GJ, Koevoets P *et al.* 2003. Screening of natural products extracts  
45  
46 for the presence of phosphodiesterase inhibitors using liquid chromatography  
47  
48 coupled online to parallel biochemical detection and chemical characterization. J  
49  
50 Biomol Screen **8**:421-429.  
51  
52  
53 Schmid I, Sattler I, Grabley S *et al.* 1999. Natural Products in High Throughput  
54  
55 Screening: Automated High-Quality Sample Preparation. J Biomol Screen **4**:15-  
56  
57 25.  
58  
59  
60

- 1  
2  
3  
4 Schobel U, Frenay M, Van Elswijk DA *et al.* 2001. High resolution screening of  
5  
6 plant natural product extracts for estrogen receptor  $\alpha$  and ER $\beta$  binding activity using  
7  
8 an online HPLC-MS biochemical detection system. *J Biomol Screen* **6**:291-303.  
9  
10  
11 Schroeder FC, Gibson DM, Churchill ACL *et al.* 2007. Differential analysis of 2D  
12  
13 NMR spectra: new natural products from a pilot-scale fungal extract library.  
14  
15 *Angew Chem Int Ed* **119**:919-922.  
16  
17  
18 Shyur L-F, Yang N-S. 2008. Metabolomics for phytomedicine research and drug  
19  
20 development. *Curr Opin Chem Biol* **12**:66-71.  
21  
22  
23 Smith A. 2002. Screening for drug discovery: The leading question. *Nature* **418**:453-  
24  
25 459.  
26  
27  
28 Sumner LW, Mendes P, Dixon RA. 2003. Plant metabolomics: large-scale  
29  
30 phytochemistry in the functional genomics era. *Phytochemistry* **62**:817-836.  
31  
32  
33 Thomas JJ, Shen Z, Crowell JE *et al.* 2001. Desorption/ionization on silicon (DIOS):  
34  
35 A diverse mass spectrometry platform for protein characterization. *Proc Natl*  
36  
37 *Acad Sci U S A* **98**:4932-4937.  
38  
39  
40 Tikunov Y, Lommen A, de Vos CHR *et al.* 2005. A novel approach for nontargeted  
41  
42 data analysis for metabolomics: Large-scale profiling of tomato fruit volatiles.  
43  
44 *Plant Physiol* **139**:1125-1137.  
45  
46  
47 Ulrich-Merzenich G, Zeitler H, Jobst D *et al.* 2007. Application of the "-Omic-"  
48  
49 technologies in phytomedicine. *Phytomedicine* **14**:70-82.  
50  
51  
52 Vaidyanathan S, Jones D, Broadhurst D *et al.* 2005. A laser desorption ionisation  
53  
54 mass spectrometry approach for high throughput metabolomics. *Metabolomics*  
55  
56 **1**:243-250.  
57  
58  
59  
60

- 1  
2  
3  
4 van der Greef J, Smilde AK. 2005. Symbiosis of chemometrics and metabolomics:  
5  
6 past, present, and future. *J Chemometr* **19**:376-386.  
7  
8  
9 van der Kooy F, Verpoorte R, Marion Meyer JJ. 2008. Metabolomic quality control  
10  
11 of claimed anti-malarial *Artemisia afra* herbal remedy and *A. afra* and *A. annua*  
12  
13 plant extracts. *South African Journal of Botany* **74**:186-189.  
14  
15  
16 van Elswijk DA, Irth H. 2002. Analytical tools for the detection and characterization  
17  
18 of biologically active compounds from nature. *Phytochemistry Reviews* **1**:427-  
19  
20 439.  
21  
22  
23 Verpoorte R. 1998. Exploration of nature's chemodiversity: the role of secondary  
24  
25 metabolites as leads in drug development. *Drug Discov Today* **3**:232-238.  
26  
27  
28 Verpoorte R, Choi Y, Kim H. 2007. NMR-based metabolomics at work in  
29  
30 phytochemistry. *Phytochem Rev* **6**:3-14.  
31  
32  
33 Verpoorte R, Choi Y, Mustafa N *et al.* 2008. Metabolomics: back to basics.  
34  
35 *Phytochem Rev* **7**:525-537.  
36  
37  
38 Verpoorte R, Choi YH, Kim HK. 2005. Ethnopharmacology and systems biology: A  
39  
40 perfect holistic match. *J Ethnopharmacol* **100**:53-56.  
41  
42  
43 Verpoorte R, Kim H, Choi Y. 2006. Plants as source for medicines: New  
44  
45 perspectives. In: Bogers R, Craker L, Lange D, editors. *Medicinal and Aromatic*  
46  
47 *Plants: Agricultural, Commercial, Ecological, Legal, Pharmacological and Social*  
48  
49 *Aspects*. Wageningen: Wageningen UR. p 261-273.  
50  
51  
52 Walters WP, Namchuk M. 2003. Designing screens: how to make your hits a hit. *Nat*  
53  
54 *Rev Drug Discov* **2**:259-266.  
55  
56  
57  
58  
59  
60

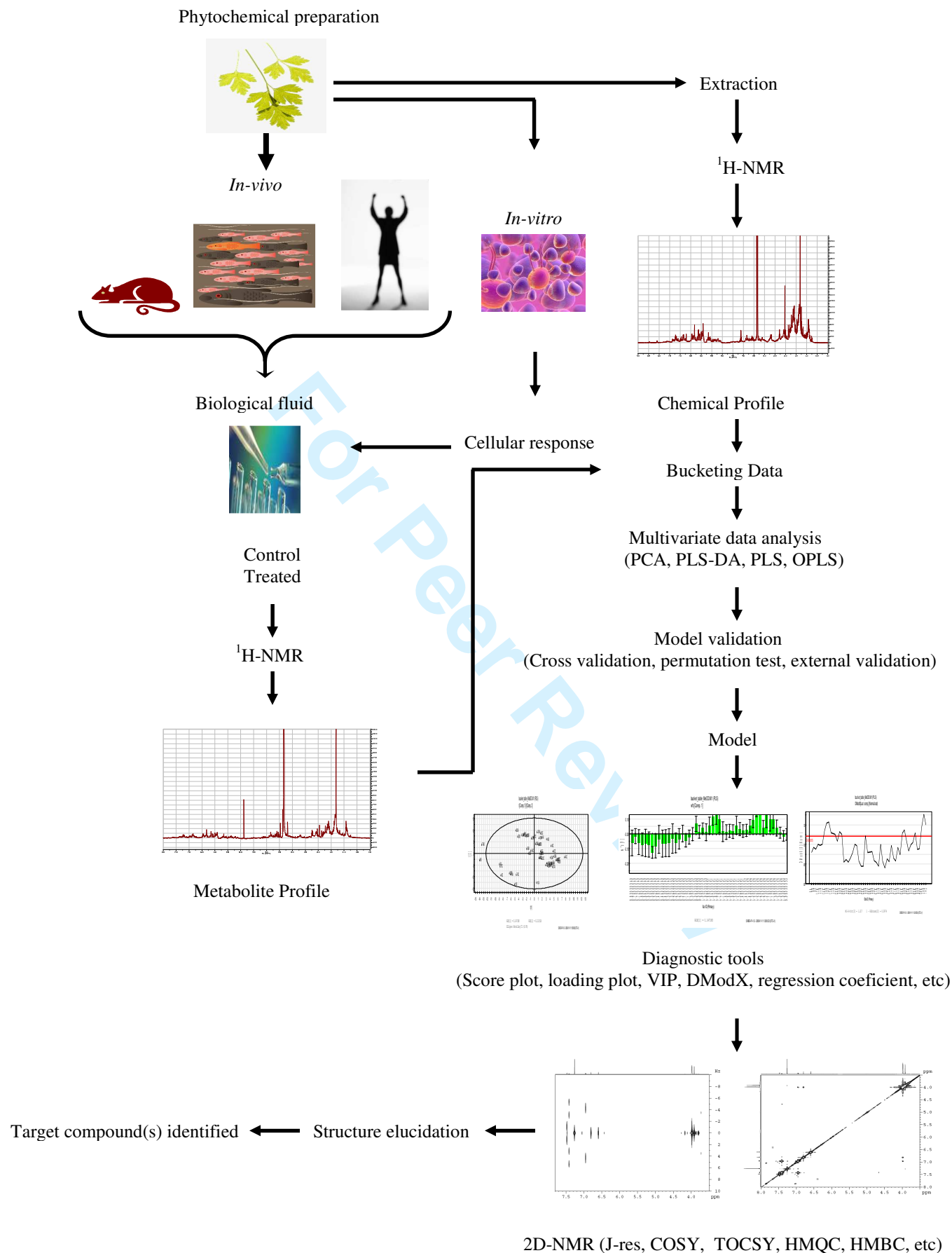
- 1  
2  
3  
4 Wang M, Lamers R-JAN, Korthout HAAJ *et al.* 2005. Metabolomics in the context  
5  
6 of systems biology: bridging traditional Chinese medicine and molecular  
7  
8 pharmacology. *Phytother Res* **19**:173-182.  
9  
10  
11 Wang P, Liang Y, Zhou N *et al.* 2007. Screening and analysis of the multiple  
12  
13 absorbed bioactive components and metabolites of Dangguibuxue decoction by  
14  
15 the metabolic fingerprinting technique and liquid chromatography/diode-array  
16  
17 detection mass spectrometry. *Rapid Commun Mass Spectrom* **21**:99-106.  
18  
19  
20  
21 Wang Y, Tang H, Nicholson JK *et al.* 2004. Metabolomic strategy for the  
22  
23 classification and quality control of phytomedicine: a case study of chamomile  
24  
25 flower (*Matricaria recutita* L.). *Planta Med* **70**:250-255.  
26  
27  
28  
29 Ward JL, Baker JM, Beale MH. 2007. Recent applications of NMR spectroscopy in  
30  
31 plant metabolomics. *FEBS J* **274**:1126-1131.  
32  
33  
34  
35 Williams G. 2004. Advances in high throughput screening. *Drug Discov Today*  
36  
37 **9**:515.  
38  
39  
40  
41 Williard X, Pop I, Bourel L *et al.* 1996. Combinatorial chemistry: a rational approach  
42  
43 to chemical diversity. *Eur J Med Chem* **31**:87-98.  
44  
45  
46  
47 Wold S, Sjöström M, Eriksson L. 2001a. PLS-regression: a basic tool of  
48  
49 chemometrics. *Chemom Intell Lab Syst* **58**:109-130.  
50  
51  
52  
53 Wold S, Trygg J, Berglund A *et al.* 2001b. Some recent developments in PLS  
54  
55 modeling. *Chemom Intell Lab Syst* **58**:131-150.  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

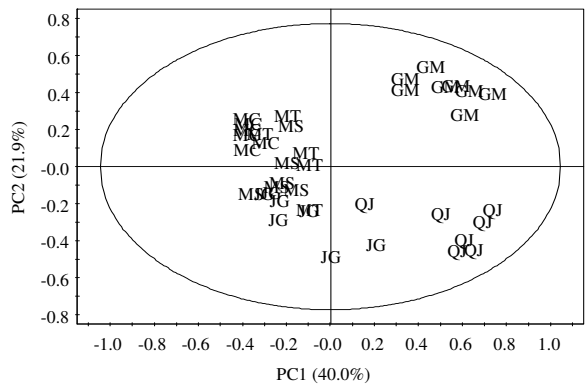
- 1  
2  
3  
4 Wu B, Yan S, Lin Z *et al.* 2008. Metabonomic study on ageing: NMR-based  
5  
6 investigation into rat urinary metabolites and the effect of the total flavone of  
7  
8 *Epimedium*. *Mol Biosyst* **4**:855-861.  
9  
10  
11 Yi L-z, Yuan D-l, Liang Y-z *et al.* 2007. Quality control and discrimination of  
12  
13 Pericarpium Citri Reticulatae and Pericarpium Citri Reticulatae Viride based on  
14  
15 high-performance liquid chromatographic fingerprints and multivariate statistical  
16  
17 analysis. *Anal Chim Acta* **588**:207-215.  
18  
19  
20 Zhang X, Wu H, Liao P *et al.* 2006. NMR-based metabonomic study on the subacute  
21  
22 toxicity of aristolochic acid in rats. *Food Chem Toxicol* **44**:1006-1014.  
23  
24  
25 Zhao X, Zhang Y, Meng X *et al.* 2008. Effect of a traditional Chinese medicine  
26  
27 preparation Xindi soft capsule on rat model of acute blood stasis: A urinary  
28  
29 metabonomics study based on liquid chromatography-mass spectrometry. *J*  
30  
31 *Chromatogr B* **873**:151-158.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

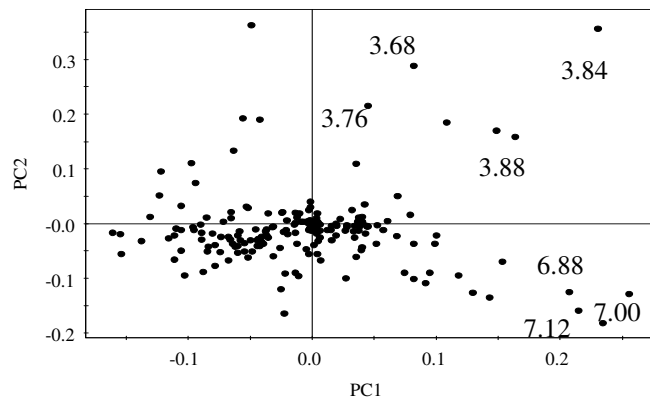
Figure 1. Schematic illustration of metabolomics and multi-variate data application in bioactivity screening of phytochemical preparation. Multi-variate data analysis permits the correlation of metabolite profile obtained from bioactivity test with chemical profile of phytochemical preparation, thus the efficacy involving either single or multi-synergistically active compounds can be viewed.

Fig. 2. A. Score plot of PCA of *Galphimia glauca* extract from six different locations. GM and QJ samples were clearly separated from others. B. Signals of the galphimine C-27 methyl ester proton at  $\delta$  3.68 – 3.88 ppm as appeared on the loading plot were the main factors for the separation C. Partial least square discriminant analysis (PLS-DA) allows further separation base on the anxiolytic activity. D. Galphimines and quercetin analogs were clearly shown as responsible compounds for the separation (adapted from Cardoso-Taketa *et al.*, 2008, with permission from Thieme Medical Publisher).

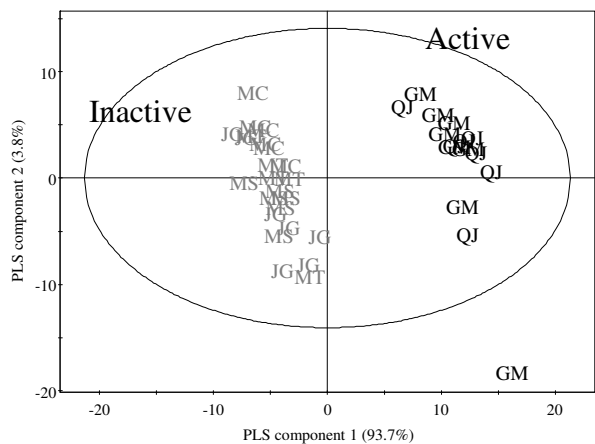




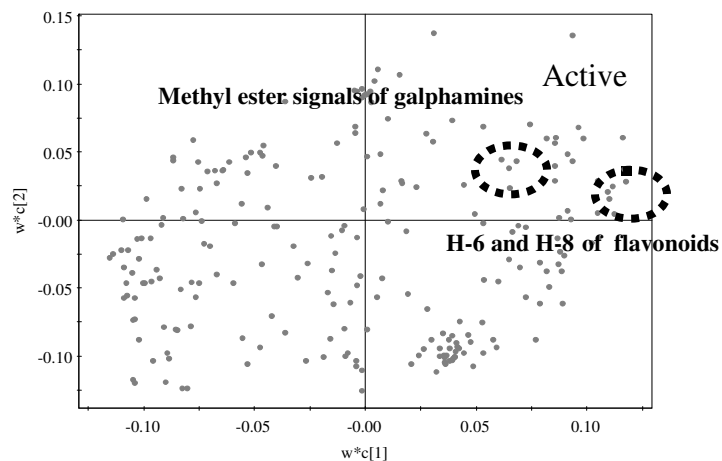
A



B



C



D