Plant Metabolomics Applications in the *Brassicaceae*: Added Value for Science and Industry

R.D. Hall[1, 2, 3, a], R.C.H. de Vos[1, 2, 3] and J.L. Ward[4]
[1]Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands
[2]Centre for BioSystems Genomics, P.O. Box 98, 6700 AA Wageningen, The Netherlands
[3]Netherlands Metabolomics Centre, Einsteinweg 55, 2333 CC Leiden, The Netherlands
[4]National Centre for Plant and Microbial Metabolomics, Rothamsted Research, West Common, Harpenden, Herts. AL5 2JQ, UK

Keywords: *Brassica*, *Arabidopsis*, metabolomics, mass spectrometry, nuclear magnetic resonance

Abstract

Crops from the family *Brassicaceae* represent a diverse and very interesting group of plants. In addition, their close relationship with the model plant, *Arabidopsis thaliana*, makes combined research on these species both scientifically valuable and of considerable commercial importance. In the post-genomics era, much effort is being placed on expanding our capacity to use advanced technologies such as proteomics and metabolomics, to broaden our knowledge of the molecular organization of plants and how genetic differences are translated into phenotypic ones. Metabolomics in particular is gaining much attention mainly due both to the comprehensiveness of the technology and also the potentially close relationship between biochemical composition (including human health-related phytochemicals) and phenotype. In this short review, a brief introduction to the main metabolomics technologies is given taking examples from research on the *Brassicaceae* for illustration.

INTRODUCTION TO THE TECHNOLOGY

Plant metabolomics is a rapidly developing technology which has the aim to provide us with exceptionally rich data on the biochemical composition of plant materials (Fig. 1). The latest extraction, separation and metabolite detection techniques are being applied and modified to this purpose. Generally, Mass Spectrometry (MS) preceded by LC (liquid chromatography) or GC (gas chromatography) for the initial separation compounds, or Nuclear Magnetic Resonance (NMR) are used for comprehensive metabolite detection (for detailed reviews see Hall, 2006; Saito et al., 2006). Such information-rich datasets require dedicated biostatistics and bioinformatics tools for data management, storage and mining, without which we cannot translate such information into biologically-relevant knowledge. The fields of application are already highly diverse and especially in the *Brassicaceae*, the topic of this short review.

In fundamental plant research, metabolomics has already gained a very strong foothold, creating new insights into the biochemical composition of plants and their molecular organization. However, metabolomics has also not gone unnoticed in the field of applied plant science. In many areas such as plant breeding, food production and food processing, the potential of metabolomics research has also been recognized as a potentially valuable approach to generate novel information of value in a commercial context (Fernie and Schauer, 2008). This information may be of direct use in defining new or improved production strategies within the food industry but also may be of direct value to the ever-concerning consumer (Hall et al., 2008). For the *Brassicaceae*, which includes the model plant Arabidopsis, the use of metabolomics approaches to help answer both fundamental and applied science questions is growing exponentially (see Keurentjes et al., 2008).

For plant metabolomics technologies, despite current limitations, any literature

[a] robert.hall@wur.nl

search will reveal a broad range of areas of application (Hall, 2006). Firstly, metabolomics is widely being used to expand our basic knowledge and understanding of key metabolic pathways and how these are influenced by genetical and environmental factors. Arabidopsis is playing a central role here (Keurentjes et al., 2006). Such knowledge will, in turn, lead to a deeper insight into the genetic mechanisms controlling these processes and hence, will enable more directed approaches for the modification of plant biochemical composition. There is great interest in exploiting such knowledge (translated into e.g. genetic markers) to develop additional tools for use, for example, in breeding programmes focused on improving the quality/nutritional value of food materials (Keurentjes et al., 2008; Fernie and Schauer, 2008). Much early work for example, focused on tomato, where a wide range of metabolomics approaches based on GCMS, LCMS and NMR have been used to define differences between fruits with genetically distinct backgrounds (Bovy et al., 2007). Through this research, we are gaining a much broader knowledge on how the metabolic composition of this crop fruit model and changes during development and how this is under the influence of genetic factors (Schauer et al., 2006, 2008). As a result, we have already a series of new genetic markers for developmental and quality traits in this commercially important fruit crop. There are also many other examples for crops such as melon, grape, tea and potato, to name but a few (Hall, 2006; Saito et al., 2006). For *Brassica*, the available information is still limited but future input via Arabidopsis will inevitably be extensive.

Metabolomics approaches are also being used to study plant volatiles which are not only of importance in flavour and fragrance phenotypes, but also have a key role to play in plant defense mechanisms (Tikunov et al., 2005). Natural volatiles function both as important signaling molecules to attract insect predators of plant pests and also as inhibitors of fungal pathogen growth (Kappers et al., 2005). GCMS analyses of such interactions are delivering the kind of information which is needed to understand these processes better and which can be used to design new approaches to enhance plant pathogen resistance. Such research is also being applied in a more ecological context (Kant and Baldwin, 2007).

Next to work on whole plants, metabolomics is also finding considerable favour in the plant/food processing industry (Hall et al., 2008). With the world's population eating more and more processed food and with the global transportation of fresh produce requiring longer storage periods, both producer and consumer are demanding a better understanding of the quality (that is, biochemical composition) of our food and how this can best be optimized for modern food processing practices (Capanoglu et al., 2008). Here also, metabolomics approaches are being used to generate the information required to identify the weak links in the food production chains and determine how these might be improved. A spin-off of this work is that metabolomics is also providing us with information on the unique biochemical profiles of food produce. This information can then also be used for tracing and tracking of food ingredients and for identifying illegal practices such as adulteration of expensive products with cheaper alternatives (Lopez-Diez et al., 2003).

The *Brassicaceae* represents a highly interesting and contrasting collection of plants. Even within *B. oleracea* alone, the diversity of vegetables is remarkable, and this species therefore represents an extremely interesting group of crops to study in the context of their biochemical composition. There is already great interest in the brassinosteroids and their contrasting profiles within the different subspecies of *Brassica* due to their proposed link with potential health-promoting properties (Malikova et al., 2007). However, next to this group of compounds, metabolomics will also reveal much more information on a wider range of metabolites which may also play key roles in other important phenotypic characteristics such as taste and disease/pest resistance. For this reason, broccoli has been chosen as the representative vegetable crop within the recently initiated EU research project META-PHOR (www.meta-phor.eu) where current technologies and bioinformatics tools are being modified and advanced to increase our ability to generate biologically-relevant information relating biochemical profile and

phenotype. Within this research both MS and NMR – based approaches are being exploited and consequently, a brief overview of the current state-of-the-art for both approaches for the *Brassicaceae* is given based upon our own results and those in the literature.

THE USE OF MASS SPECTROMETRY – BASED APPROACHES FOR METABOLOMICS OF THE *BRASSICACEAE*

For Mass Spectrometry (MS), a specific electron voltage is used at the source of the instrument to produce charged molecules (in the form of ions or free radicals) from individual metabolites. These charged molecules are subsequently channeled, filtered and counted by a sensitive detector. For MS-based metabolomics applications, crude metabolite extracts are either directly injected into the mass detector (so-called direct infusion MS) or are initially separated, on-line, through the use of gas chromatography (GC) of liquid chromatography (LC).

MS preceded by GC (GC-MS) mostly involves so-called electron impact ionization, by applying a high voltage (usually 70eV), to produce charged molecules. GC-MS approaches produce highly reproducible separation and fragmentation patterns of metabolites, which enables the development of common GC-MS based metabolite libraries (Schauer et al., 2005). GC-MS is nowadays frequently used for metabolic profiling of both volatile and nonvolatile compounds in (plant) extracts and GC coupled to a very fast-scanning mass detector, i.e. time-of-flight (TOF) MS, was in fact the first approach used for large-scale plant metabolomics (Fiehn et al., 2000). The GC-TOF MS approach uses extracts which are chemically derivatised prior to injection and covers a large variety of non-volatile metabolites. This includes mainly metabolites involved in primary metabolism, including organic and amino acids, sugars, sugar alcohols, phosphorylated intermediates (in the polar fraction of extracts), as well as lipophilic compounds such as fatty acids and sterols (in the apolar fraction). Detailed protocols for sample extraction, derivatization, analysis and subsequent data analyses are now widely available (Saito et al., 2006; Hardy and Hall, in preparation). As most primary metabolites have commercially available standard compounds, GC-MS can produce quantitative data for tens to up to a few hundred compounds involved in central metabolism. Figure 2 shows examples of GC-TOF MS chromatograms obtained using polar leaf extracts of *Brassica oleracea*. In a screening of 24 different *Brassica oleracea* genotypes, more than 1,000 relevant mass signals representing 63 different metabolites (exceeding 50 times the background level) could be detected, using the dedicated Metalign software for mass signal extraction and alignment (www.metalign.nl; free download).

A second approach using GC-MS is the profiling of (naturally) volatile compounds, e.g. from flowers like Petunia (Verdonk et al., 2003) or fruits of crop plants such as strawberry (Aharoni et al., 2000) and tomato (Tikunov et al., 2005). For example, using solid phase microextraction (SPME) coupled to a GC-MS based untargeted metabolomics approach, a total of 322 different volatiles were detected in the headspace of tomato fruits (Tikunov et al., 2005). Using correlation analysis tools, such volatile profiles can then be linked to fragrance or taste characteristics, in order to identify volatile compounds and their respective biosynthetic pathways which are responsible for the differences in these economically important quality aspects.

For analyzing semi-polar metabolites, including the large group of plant secondary metabolites such as glucosinolates, phenylpropanoids, flavonoids, alkaloids, saponins, phenolic acids and all derivatives thereof, the preferred method is LC-MS. Semi-polar metabolites can be effectively extracted using aqueous alcohol solutions and directly analyzed without derivatization (Moco et al., 2006; De Vos et al., 2007). In contrast to GC-MS, LC-MS mostly employs a soft-ionization technique such as electrospray ionization (ESI) or atmospheric pressure chemical ionization (APCI), resulting in protonated (in positive mode) or deprotonated (in negative mode) molecules. Modern high mass resolution detectors, like time-of-flight (TOF)-MS, Orbitrap-MS and Fourier Transform-Ion Cyclotron Resonance-MS (FTMS) instruments, enable the calculation of

the elemental formula of detected ions. Based on this feature of MS instruments, rapid direct-infusion MS approaches without any prior compound separation have been developed to compare rapidly plants for differences in their metabolite fingerprints (Aharoni et al., 2002; Hirai et al., 2005). However, with crude (plant) extracts such direct infusion approaches may suffer from significant artifact formation and ion suppression resulting in loss of sensitivity. Moreover, by definition, isomers (which have identical masses) cannot be discriminated without chromatographic separation.

LC preceding MS not only enables the separation and detection of isomeric compounds, which are often abundantly present in plants, but can also provide valuable structural information of the metabolites, on-line, e.g. through MS/MS fragmentation and recording absorbance spectra using photodiode array (PDA) detection. High chromatographic resolution, such as Ultra-Performance LC (UPLC), in combination with high mass resolution, such as TOF-MS or FTMS, may enable the detection of up to several hundreds of compounds in a single crude plant extract (Moco et al., 2006; Böttcher et al., 2008; Ijima et al., 2008). Such comprehensive metabolite analysis techniques have been used to determine the exact differences in metabolite composition between transgenic or mutant and control plants, thereby enabling the functional analysis of genes or mutations (Bino et al., 2005, Beekwilder et al., 2008). LC-MS based metabolomics approaches have also been used in a large-scale metabolomics screening of an *Arabidopsis* recombinant inbred line (RIL) population (Keurentjes et al., 2006). Here the variation in metabolic profiles was linked to genetic variation, in order to detect quantitative trait loci (QTL's) regulating metabolite composition. This so-called genetical metabolomics approach enabled the identification of both known and novel QTL's, including loci regulating the production of health-related compounds such as glucosinolates and flavonoids. With regard to *Brassica* crop plants, within several projects we are currently screening metabolite variation using LC-MS based metabolomics in a range of *B. rapa* and *B. oleracea* genotypes and crossing populations, in order to link variation in metabolite profiles to phenotypic traits and markers suitable for breeding applications. For example, in an LC-QTOF MS screening of 170 *B. rapa* genotypes, more than 22000 unique mass signals, representing more than 600 different metabolites (exceeding 50 times background level), were detected. Within the EU-META-PHOR project, the focus with regard to *Brassicaceae* is on broccoli, and the effects of variation due to cultivar, growing conditions and processing differences. Figure 3 shows a typical example of an LC-MS profile from a freshly harvested broccoli flower head. Using Metalign to extract and align all mass signals and subsequent multivariate analyses of the resulting LCMS data set, 3 different broccoli cultivars grown in the field could readily be separated based on differences in secondary metabolites extracted from their flower heads (Fig. 4).

THE USE OF NMR – BASED APPROACHES FOR METABOLOMICS OF THE *BRASSICACEAE*

^{1}H-NMR fingerprinting of plant extracts has found widespread use for the rapid comparison of lines and is particularly suitable for classifying cultivars, ecotypes and mutants and also has found use in the study of substantial equivalence of GM versus non-GM material (Ward et al., 2003, 2007; Baker et al., 2006). The technique is well suited to automation and thus is capable of producing comparative fingerprints from several thousands of samples in a single experiment. Most research groups use a polar solvent extract, usually water/methanol mixtures, and this gives rise to NMR spectra containing overlapping signals for metabolites such as carbohydrates, amino-acids, organic acids, nucleotides and nucleosides, phenylpropanoids, flavonoid glycosides, and in the *Brassicaceae*, also intact glucosinolates (Fig. 5). These fingerprints are extremely reproducible and can be readily aligned and then compared using Principal Components Analysis (PCA) to discover similarities and differences between plant lines. These can be assigned to particular metabolites using libraries of spectra of pure compounds as a reference. In the META-PHOR project, we are developing software that is designed to

automate the interpretation of loading plots by computerised searching and matching against metabolite libraries.

In order to carry out larger studies in metabolomics it is necessary to often include multiple lines, treatments or time points as well as the usual biological and analytical replicates. This can make the total number of samples, in one experiment, very large and for some studies (e.g. time courses of perturbation effects, QTL studies) can reach many hundreds or thousands. It is imperative that a robust technique is used to collect the analytical data in order to minimize the amount of variance due to analytical variation in the final dataset. The EU-META-PHOR project has designed a range of intra and inter-lab studies to determine the reproducibility of analytical techniques and methodologies and aims to assess the robustness of different techniques used in the field of metabolomics. In this vein, the NMR "ring experiments", using broccoli as the test material have demonstrated the superb stability of the NMR technique and has suggested that data collected at different sites on instruments of different field strength can indeed be modelled together and yield the same biological result (Ward et al., unpublished). This important finding suggests great potential for multi laboratory collaborations towards collective data repositories of ^{1}H NMR data.

A comprehensive review on health promoting metabolites in *Brassica* vegetables (Podsedek, 2007) describes the major detrimental effects of tissue storage and post-harvest processing. Within the META-PHOR project the effect of post harvest processing of plant tissue and the differences due to grower location and local practices have also been assessed. Using three broccoli cultivars (Monaco, Chevalier and Iron Man) and four producer locations we have been able to discern differences due to cultivar irrespective of location or date of harvest (Fig. 6). Further mining of the NMR dataset allowed us to discern metabolite differences based on producer location and date of harvest for the Monaco samples (data not shown). The ability to track changes in primary and secondary metabolites caused directly by location, harvest time or tissue processing is of great importance to commercial growers. As many of the metabolites seen in a typical ^{1}H NMR spectrum of *Brassica* material are natural antioxidants or health beneficial compounds, the losses due to processing in particular need to be taken into account especially when calculating the dietary intake of e.g. antioxidants from processed food.

^{1}H NMR has also been used, within the META-PHOR project to assess the similarities and differences in metabolite profile across a range of brassicas (Table 1). Analysis of the profiles was carried out by a visual inspection, by Principal Component Analysis and by KNN clustering (Fig. 7). By far the most diverse samples within this set were identified as mustard, swiss chard, Chinese cabbage and swede. These samples separated by both PCA methods and formed a separate grouping in the KNN clustering analysis. Other interesting findings were that broccoli samples separated depending on the colour of the variety, with the purple varieties clustering in a separate group with other red/purple brassicas (e.g. Brussel sprout Falstaff and kohl rabi Delicacy Purple). Despite the profiles of the different *B. Oleracea* cabbage samples looking (qualitatively) very similar to each other, clear differences could be seen in the intensities of key primary metabolites (carbohydrates and amino acids) and this caused a separation of these samples into a number of different groups. This was especially true of the different varieties of Savoy cabbage in the study which clustered separately with other members of the *Brassica* family.

In terms of using ^{1}H NMR to study the effects of perturbation (chemical or environmental), the technique has already been used to improve the effect in *Brassica rapa* leaves treated with methyl jasmonate (Liang et al., 2006). This study demonstrated the use of two-dimensional NMR for the identification of phenylpropanoids in *B. rapa* leaves whose levels increased when exposed to the jasmonic acid derivative implying a defence role for the hydroxycinnamate family of metabolites. The structural elucidation elements of this study demonstrate the power of ^{1}H NMR in natural product structure determination. Metabolomics experiments, whilst generating extremely large datasets yielding clues to decipher modifications in the biochemical machinery of the tissue under

study, often reach the point of needing to identify a high number of "unknowns". It is this final step in any metabolomics experiment that is both the most difficult and the most time consuming. The ability of the NMR technique to capture information on both ^{1}H and ^{13}C nuclei often provide the necessary information to assign a chemical structure to discriminatory peaks arising from these complex datasets. Whilst other techniques such as LC-MS and GC-MS are employed by many laboratories for the screening or metabolite fingerprinting approach, it is often NMR which is required for final structural confirmation should an authentic standard not be available. This approach of employing a variety of techniques is common and has been employed with success in the *Brassica* metabolomics community. Characterization of a complex family of sinapate esters from *B. napus* seeds was recently achieved by utilizing a combined LC-MS/MS approach in conjunction with NMR for final structural identification of fifteen seed constituents including phenylpropanoids and flavonoids (Baumert et al., 2005). Similarly, for benzoic acid and indole glucosinolates in *Arabidopsis thaliana* NMR has played a key role in structural identification (Reichelt et al., 2002; Agerbirk et al., 2001) of, in these cases novel glucosinolates.

Much of the development work in plant NMR fingerprinting has been carried out in *Arabidopsis* (e.g. Ward et al., 2003) and has now reached a level of automation such that robotics are being used to weigh plant powders, prepare the solvent extracts and dispense to solvent to the appropriate vessels for analysis. With the ever-increasing sample numbers and complexity in experimental design, these systems will provide the capability to provide comprehensive coverage of a range of biological questions within the same biological experiment.

PROSPECTS AND CONCLUSIONS

The fields of application for metabolomics technologies are growing rapidly. Current reports on *Brassica* are not yet extensive but research using the plant model species *Arabidopsis thaliana* is already considerable and is expanding significantly. While clearly all species and genotypes are basically unique regarding their individual metabolic profiles and work on each species/subspecies in *Brassica* needs to be performed before a true understanding of their biochemical composition is possible, *Brassica* researchers can still benefit hugely from the information generated using *Arabidopsis*. Some of the key limitations and bottlenecks to metabolomics approaches, such as metabolite identification and data management strategies, are primarily being tackled using Arabidopsis and the data and information generated is directly translatable to commercial *Brassica* genotypes. Hence, effective interaction and information exchange between the more fundamental *Arabidopsis* researchers and those focusing more on applied research into the crop species will be greatly beneficial.

ACKNOWLEDGEMENTS

This work was financed by the EU Framework 6 programme project MET-PHOR (2006-FOOD-CT-036220). RDH and RdeV acknowledge additional financing form the Centre for Biosystems Genomics and The Netherlands Metabolomics Centre, both initiatives under the auspices of the Netherlands Genomics Initiative. JW acknowledges further support form the Biotechnology and Biological Sciences Research Council for funding to set up the Rothamsted Research Metabolomics Centre (MeT-RO).

Literature Cited

Agerbirk, N., Petersen, B.L., Olsen, C.E., Halkier, B.A. and Nielsen, J.K. 2001. 1,4-Dimethoxyglucobrassicin in *Barbarea* and 4-hydroxyglucobrassicin in *Arabidopsis* and *Brassica*. 49:1502–1507.

Aharoni, A., Keizer, L.C.P., Bouwmeester, H.J., Sun, Z., Alvarez, H.M., Verhoeven, H.A., Blaas, J., van Houwelingen, A.M.M.L., De Vos, R.C.H., van der Voet, H., Jansen, R.C., Guis, M., Mol, J., Davis, R.W., Schena, M., van Tunen, A.J. and O'-Connell, A.P. 2000. Identification of the SAAT gene involved in strawberry flavor

biogenesis by use of DNA microarrays. Plant Cell 12:647–661.
Aharoni, A., De Vos, R.C.H., Verhoeven, H.A., Maliepaard, C.A., Kruppa, G., Bino, R. and Goodenowe, D.B. 2002. Nontargeted metabolome analysis by use of Fourier Transform Ion Cyclotron Mass Spectrometry. OMICS. p.217–34.
Baker, J.M., Hawkins, N.D., Ward, J.L., Lovegrove, A., Napier, J.A., Shewry, P.R. and Beale, M.H. 2006. A metabolomic study of substantial equivalence of field-grown GM wheat. Plant Biotechnology J. 4:381–392.
Baumert, A., Milkowski, C., Schmidt, J., Nimtz, M., Wray, V. and Strack, D. 2005. Formation of a complex pattern of sinapate esters in *Brassica napus* seeds, catalysed by enzymes of a serine carboxypeptidase-like acyltransferase family? Phytochemistry 66:1334–1345.
Beekwilder, J., Van Leeuwen, W., Van Dam, N.M., Bertossi, M., Grandi, V., Mizzi, L., Soloviev, M., Szabados, L., Molthoff, J.W., Schipper, B., Verbocht, H., De Vos, R.C.H., Morandini, P., Aarts, M.G.M. and Bovy, A.G. 2008. The impact of the absence of aliphatic glucosinolates on insect herbivory in *Arabidopsis*. Plos ONE 3, e2068.
Bino, R.J., De Vos, R.C.H., Lieberman, M., Hall, R.D., Bovy, A.G., Jonker, H.H., Tikunov, Y., Lommen, A., Moco, S.I.A. and Levin, I. 2005. The light-hyperresponsive high pigment-2dg mutation of tomato: alterations in the fruit metabolome. New Phytologist 166:427–438.
Böttcher, C., von Roepenack-Lahaye, E., Schmidt, J., Schmotz, C., Neumann, S., Schell, D. and Clemens, S. 2008. Metabolome analysis od biosynthetic mutants reveals a diversity of metabolic changes and allow identification of a large number of new compounds in *Arabidopsis*. Plant Physiol. 147:2107–2120.
Bovy, A., Schijlen, E.G.W.M. and Hall, R.D. 2007. Metabolic engineering of flavonoids in tomato (*Solanum lycopersicum*): the potential for metabolomics. Metabolomics 3:399–412.
Capanoglu, E., Beekwilder, J., Boyacioglu, D., Hall, R.D. and Vos, C.H.R. 2008. Changes in antioxidants and metabolite profiles during production of tomato paste. J. of Agricultural Chemistry 56:964–973.
De Vos, R.C.H., Moco, S., Lommen, A., Keurentjes, J.J., Bino, R.J. and Hall, R.D. 2007 Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nature Protocols 2:778–791.
Fernie, A.R. and Schauer, N. 2008. Metabolomics – assisted breeding: a viable option for crop improvement. Trends in Genetics, doi:10.1016/j.tig.2008.10.010.
Fiehn, O., Kopka, J., Dörmann, P., Altmann, T., Trethewey, R.N. and Willmitzer, L. 2000. Metabolite profiling for plant functional genomics. Nature Biotechnology 18:1157–1161.
Hall, R.D. 2006. Plant metabolomics: from holistic hope to hype to hot topic. New Phytologist Tansley Review, New Phytologist 169:453–468.
Hall, R.D., Brouwer, I. and Fitzgerald, M.E. 2008. Plant metabolomics and its potential for human nutrition. Physiologia Plantarum 132:162–175.
Hardy, N. and Hall, R.D. (Eds.) 2009. Methods in Molecular Biology: Plant Metabolomics (in preparation).
Hirai, M.Y., Klein, M., Fujikawa, Y., Yano, M., Goodenowe, D.B., Yamazaki, Y., Kanaya, S., Nakamura, Y., Kitayama, M., Suzuki, H., Sakurai, N., Shibata, D., Tokuhisa, J., Reichelt, M., Gershenzon, J., Papenbrock, J. and Saito, K. 2005. Elucidation of gene-to-gene and metabolite-to-gene networks in *Arabidopsis* by integration of metabolomics and transcriptomics. J. Biol. Chem. 280:25590–25595
Iijima, Y., Nakamura, Y., Ogata, Y., Tanaka, K., Sakurai, N., Suda. K., Suzuki, T., Suzuki., H., Okazaki, K., Kitayama, M., Kanaya, S., Aoki, K. and Shibata., D. 2008. Metabolite annotations based on the integration of mass spectral information. Plant J. 54:949–962.
Kant, M.R. and Baldwin, I.T. 2007. The ecogenetics and ecogenomics of plant-herbivore interactions: rapid progress on a slippery road. Current Opinion in Genetics and

Development 17:519–524.
Kappers, I.F., Aharoni, A., van Herpen, T.W., Luckerhoff, L.L., Dicke, M. and Bouwmeester, H.J. 2005. Genetic engineering of terpenoids attracts bodyguards to *Arabidopsis*. Science 309:2070–2072.
Keurentjes, J.J., Koornneef, M. and Vreugdenhil, D. 2008. Quantitative genetics in the age of omics. Current Opinion in Plant Biology 11:123–128.
Keurentjes, J.J.B., Fu, Y., De Vos, C.H.R., Lommen, A., Hall, R.D., Bino, R.J., Van Der Plas, L.H.W., Jansen R.C., Vreugdenhil, D. and Koornneef, M. 2006. The genetics of plant metabolism. Nature Genetics 38:842–849.
Liang, Y.-S., Kim, H.K., Lefeber, A.W.M., Erkelens, C., Choi, Y.H. and Verpoorte, R. 2006. Identification of phenylpropanoids in methyl jasmonate treated *Brassica rapa* leaves using two –dimensional nuclear magnetic resonance. J. Chrom. A. 1112:148–155.
Lopez-Diez, E.C., Bianchi, E.G. and Goodacre, R. 2003. Rapid quantitative assessment of the adulteration of virgin olive oils with hazelnut oils using Raman spectroscopy and chemometrics. J. Agric. Food Chem. 51:6145–6150.
Malikova, J., Swaczynova, J., Kolar, Z. and Strnad, M. 2007. Anticancer and anti-proliferative activity of natural Brassinosteroids. Phytochemistry 69:418–426.
Moco, S., Bino, R.J., Vorst, O., Verhoeven, H.A., de Groot, J., van Beek, T.A., Vervoort, J. and De Vos, C.H.R. 2006. A liquid chromatography-mass spectrometry-based metabolome database for tomato. Plant Physiol. 141:1205–18.
Podsedek, A. 2007. Natural antioxidants and antioxidant capacity of *Brassica* vegetables: A review. LWT Food Science and Technology 40, 1, 1–11.
Reichelt, M., Brown, P.D., Schneider, B, Oldham, N.J., Stauber, E., Tokuhisa, J., Kliebenstein, D.J., Mitchell-Olds, T. and Gershenzon, J. 2002. Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*. Phytochemistry 59:663–671.
Saito, K., Dixon, R. and Willmitzer, L. (Eds.) 2006. Plant Metabolomics. Springer-Verlag, Heidelberg, Germany.
Schauer, N., Steinhauser, D., Strelkov, S., Schomburg, D., Allison, G., Moritz, T., Lundgren, K., Roessner-Tunali, U., Forbes, M.G., Willmitzer, L., Fernie, A.R. and Kopka, J. 2005. GC-MS libraries for the rapid identification of metabolites in complex biological samples. Febs Letters 579:1332–1337.
Schauer, N., Semel, Y., Balbo, I., Steinfath, M., Repsilber, B., Selbig, J., Pleban, T., Zamir, D. and Fernie, A.R. 2008. Mode of inheritance of primary metabolic traits in tomato. Plant Cell 20:509–523.
Schauer, N., Semel, Y., Roessner, U., Gur, A., Balbo, I., Cararri, F., Plaban, T., Perez-Melis, A., Bruedigam, C., Kopka, J., Willmitzer, L., Zamir, D. and Fernie, A.R. 2006. Comprehensive metabolic profiling and phenotypingof interspecific introgression lines for tomato improvement. Nature Biotechnology 24:447–454.
Verdonk, J.C., De Vos R.C.H., Verhoeven, H.A., Haring, M.A., Tunen, A.J. van and Schuurink, R.C. 2003. Regulation of floral scent production in petunia revealed by targeted metabolomics. Phytochemistry 62:997–1008.
Tikunov, Y., Lommen, A., De Vos, C.H.R., Verhoeven, H.A., Bino, R.J., Hall, R.D., Lindhout, P. and Bovy, A.G. 2005. A novel approach for non-targeted analysis for metabolomics: large scale profiling of tomato fruit volatiles. Plant Physiol. 139:1125–1137.
Ward, J.L., Baker, J.M. and Beale, M.H. Recent applications of NMR spectropscopy in plant metabolomics 2007. FEBS J. 274:1126–1131.
Ward, J.L., Harris, C., Lewis, J. and Beale, M.H. 2003. Assessment of ^{1}H-NMR spectroscopy and multivariate analysis as a technique for metabolite fingerprinting of *Arabidopsis thaliana*. Phytochemistry 62:949–957.

Tables

Table 1. *Brassica* samples studied using ^{1}H NMR in the diversity screen.

Vegetable	Species	Variety
Broccoli	*B. oleracea*	White Eye
Broccoli	*B. oleracea*	Late Purple Sprouting
Broccoli	*B. oleracea*	Crown and Sceptre
Cauliflower	*B. oleracea*	all year round
Brussel sprouts	*B. oleracea*	Bedford
Brussel sprouts	*B. oleracea*	Falstaff (red)
Cabbage	*B. oleracea*	Savoy King
Cabbage	*B. oleracea*	Savoy January King
Cabbage	*B. oleracea*	Greyhound
Cabbage	*B. oleracea*	Savoy Ormskirk late
Cabbage	*B. oleracea*	Wintergreen
Cabbage	*B. oleracea*	Kalibos
Kohl rabi	*B. oleracea*	Delicacy White
Kohl rabi	*B. oleracea*	Delicacy Purple
Chinese kale	*B. oleracea*	Kailaan
Borecole	*B. oleracea*	Unknown
Swiss chard	*B. vulgaris*	Lucullus
Rocket	*Eruca vesicaria*	Unknown
Rape	*B. napus*	Spring Twyford
Mustard	*Sinapis alba*	White
Swede	*B. napobrassica*	Magres
Chinese cabbage	*B. rapa*	Wong Bok
Pak choi	*B. rapa*	Unknown
Pak choi	*B. rapa*	Cotton Dwarf

Figures

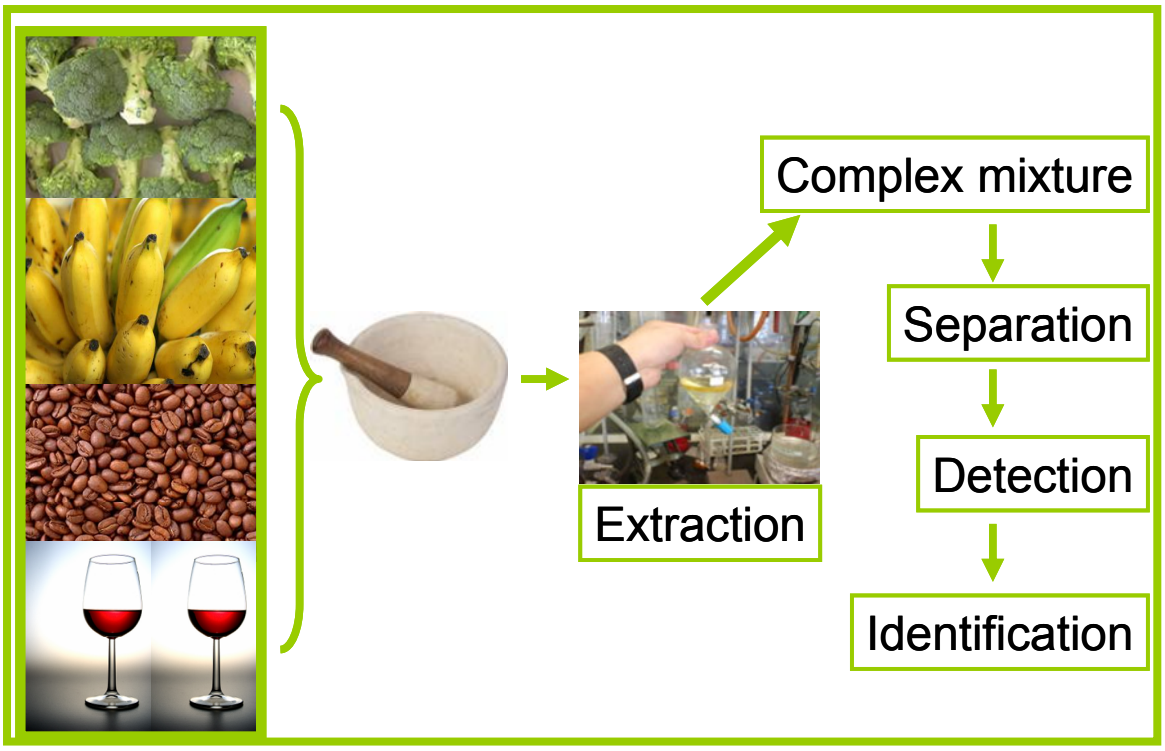


Fig. 1. The standard approach for metabolomics analysis involves optimized procedures for the extraction of complex mixtures of metabolites from any kind of plant product, using an appropriate solvent after which the mixture is usually chromatographically separated using either a liquid or gas phase after which the individual metabolites are detected using Mass Spectrometry or Nuclear Magnetic Resonance. The latter technology can also be used without any prior separation.

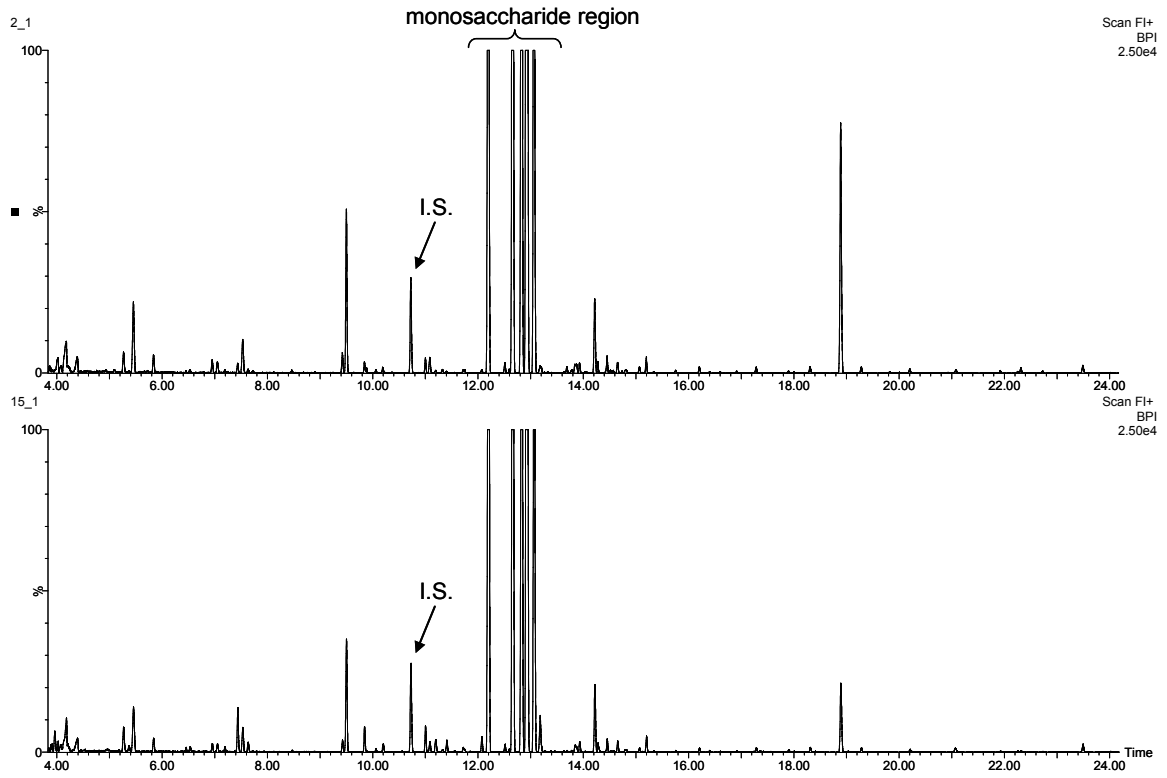
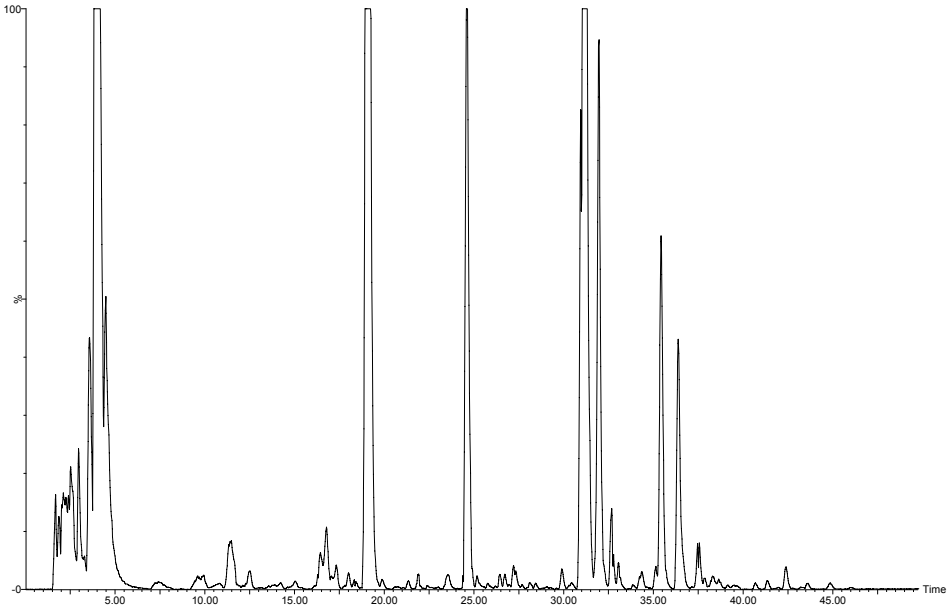


Fig. 2. GC-TOF MS profiles (similar scale) of derivatized crude polar extracts from leaves of 2 *Brassica oleracea* cultivars analyzed within a series of 24 different genotypes. The five large peaks between 12 and 13.5 min represent monosaccharides including fructose and glucose, while the differential peak at Retention Time = 18.9 min corresponds to sucrose. I.S.: internal standard (ribitol). Automated mass signal extraction and alignment using Metalign software (www.metalign.nl) and subsequent data filtering for metabolite signal redundancy (Tikunov et al., 2005) resulted in a metabolomics dataset containing relative levels of 86 polar metabolites in each genotype.



Retention time (min)

Fig. 3. A typical LC-MS chromatogram (recorded in negative ion mode) of a broccoli extract (*Brassica oleracea* 'Iron Man'). The largest peaks represent glucosinolates and flavonoids, which are highly abundant in most *Brassica* species. One hundred milligram of frozen and ground tissue (flower head) was extracted in 500 µl of 75% methanol in water, and 5 µl was injected into the LC-MS and separated on a C_{18}-reversed phase column in a 45 min period using a water-acetonitrile gradient.

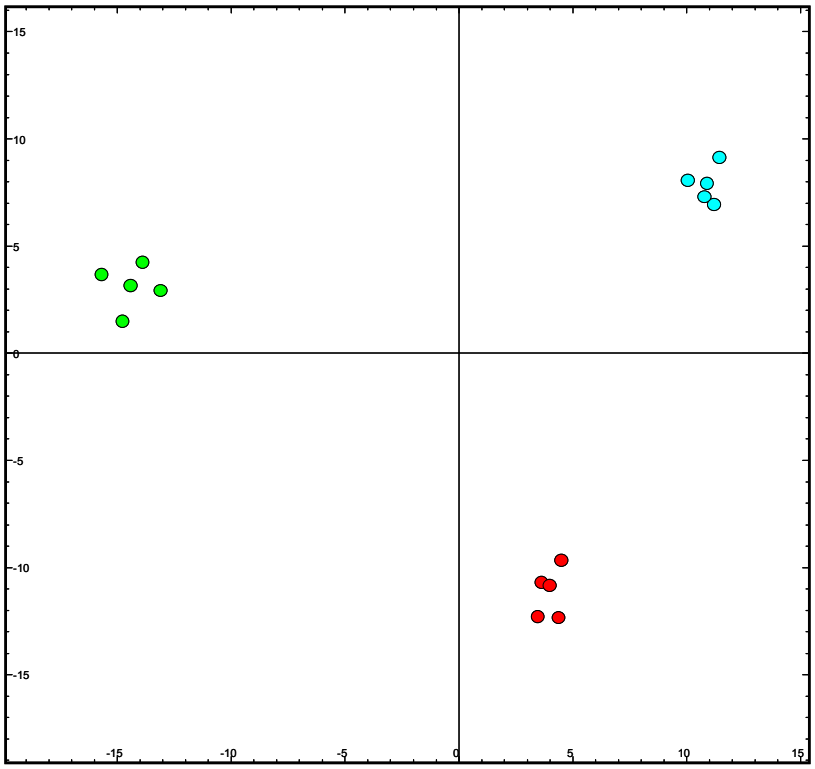


Fig. 4. Principal components analysis (PCA) of LCMS metabolomics data of aqueous-methanol extracts from the flower heads clearly separate the 3 different broccoli cultivars used: top left dots: ‘Chevalier’; top right dots: ‘Iron Man’; lower right dots: ‘Monaco’. Each dot represents an individual biological replicate; 5 replicates per sample were measured. PC1 (X axis) and PC2 (Y axis) explain 46 and 26% of the total variation, respectively.

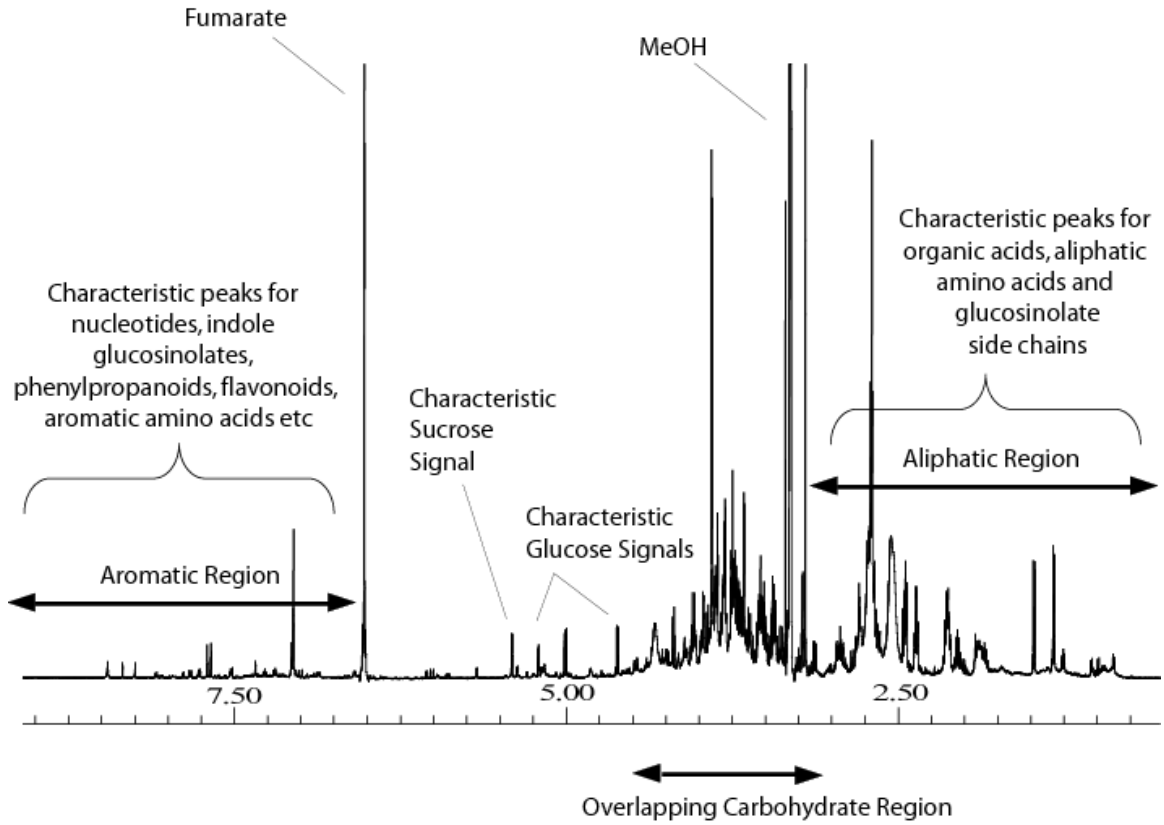


Fig. 5. 600 MHz ^{1}H NMR spectrum of polar solvent extract of *Arabidopsis*. The spectrum was obtained via extraction of freeze-dried *Arabidopsis* (Col-0, 15 mg) leaves with 80:20 D_2O:CD_3OD at 50°C for 10 min.

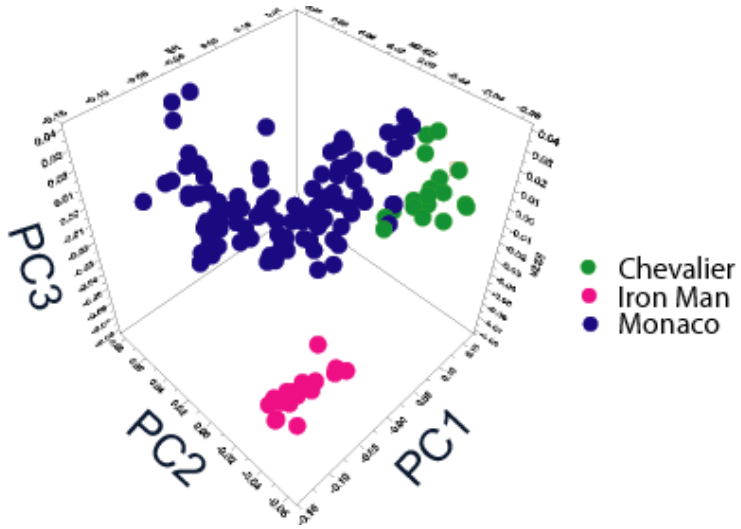


Fig. 6. 3-Dimensional PCA analysis of broccoli cultivars showing separation by cultivar irrespective of site or date of harvest following ^{1}H NMR analysis.

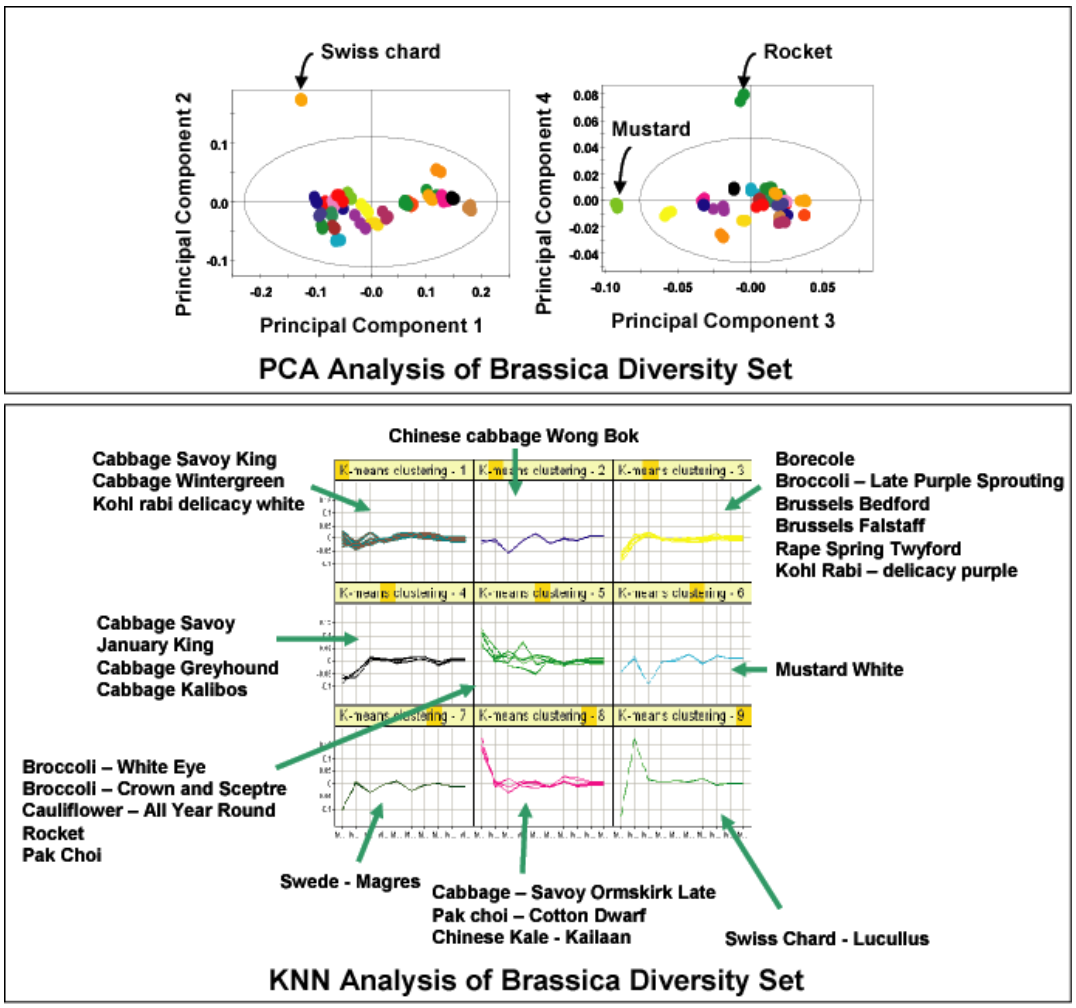


Fig. 7. Principal Components Analysis (PCA) [upper] and K Nearest Neighbour (KNN) clustering analysis [lower] of ^{1}H NMR data obtained from screening 24 different *Brassica* species (listed in lower figure). KNN cluster analysis was carried out using individual PC scores from the initial Principal Component Analysis. Clustering demonstrates the clear differences in mustard, Swiss chard, swede and Chinese cabbage and similarities between some of the other more closely-related variants.