Bulletin UASVM Agriculture, 68(2)/2011
Print ISSN 1843-5246; Electronic ISSN 1843-5386

Metabolomics-Based Systems Biology Needs Chemometrics – Former Eexperience and Case Studies

Carmen SOCACIU, Florinela FETEA, Floricuța RANGA, Raluca POP, Simona ZAVOI

University of Agricultural Sciences and Veterinary Medicine, Faculty of Agriculture, Department of Chemistry and Biochemistry, 3-5 Mănăştur Street, 400372, Cluj-Napoca, Romania. email: carmen.socaciu@usamvcluj.ro * corresponding author

Abstract. Considering the need of symbiosis between metabolomics& chemometrics in developing systems biology, we review here some relevant findings of our previous experience in performing analytical methods for metabolomic fingerprinting of fruits and food products, coupled with chemometrics as an integrated , added-value technology for systems biology. Some specific case-studies relevant for plant and food metabolomics are presented: seabuckthorn fruits and leaves, lipophilic and hydrophilic extracts or juices, aronia, black currant and bilberries. We proposed specific metabolomic-metabonomic evaluations (fingerprint and quantification) integrated by a four-steps analysis: UV-VIS spectroscopy (1), Infrared spectrometry (2), GC or HPLC ± FID, PDA or MS detection (3) and chemometry (4). By our specific case-studies we demonstrated here that the determination of biochemical markers for a specific plant or food is just a beginning of a metabolic approach. Only combining the chemical informations given by sophisticated or rapid, simple techniques associated with chemometrics' informations *via* pattern recognition (fingerprinting), data calibration and quantitative measuring, as well clustering of significant groups of samples based on their principal components are necessary to enrich the metabolic profile and approach the intergrated view of systems biology.

Keywords: Systems Biology, Metabolomics, Principal Component Analysis, Cluster Analysis, plant taxonomy , food authenticity

INTRODUCTION

I. Plant and food Metabolomics/metabonomics

Metabolomics is a new-emerging technology, developing tremendously the last decade, as part of the "omics" family which complement the large-scale gene transcript analysis (transcriptomics) and proteins fingerprint (proteomics), explaining and identifying the differences between sets of organisms (e.g. differences in genotypes and phenotypes and their classification named Chemotaxonomy) and elucidate environmental factors that influence biochemical events (Fiehn, 2002; Sumner et al., 2003; Bender, 2005; Dunn& Ellis, 2005; Lindon et al., 2007,).

Metabolomics is defined as a systematic study of (bio)chemical fingerprint which realize a metabolite profiling (small molecules) in a specific matrix (plant, food, animal or human tissue or cells) . **Metabonomics** includes quantitative measurements to identify a specific metabolic response (by key-molecules, e.g. pigments, volatiles, sterols, vitamins) **Metabolites** are end products of gene expression and enzymatic activities and reflect the activity of a certain metabolic pathway or chain in a static or dynamic manner (fluxomics). The metabolome represents the complement of all metabolites expressed in a cell, tissue or organism. A suggestive representation of the importance of metabolomics as a synergic combination of biochemistry and analytical/organic chemistry is presented in Fig.1., as well

the factors which influence the metabolic state of plants/animals/humans and food metabolome.

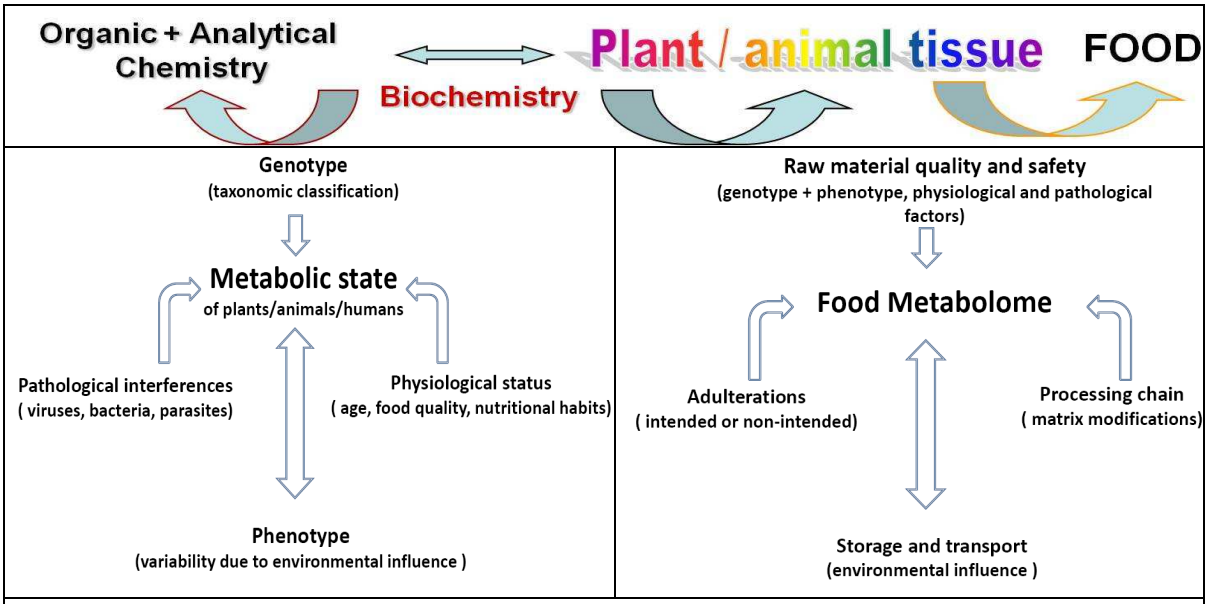

Fig.1. Metabolomics as a synergic combination of biochemistry and analytical/organic chemistry. The main factors which influence the metabolic state of plants/animals/humans and food metabolome are also represented.

II. Metabolomics need Advanced Analytics and Chemometrics

The progress registered in Analytical Chemistry was a sine-qua-non condition for metabolomics development. To mention the most important techniques which are currently used to evaluate qualitatively (fingerprinting) or quantitatively (metabolic profiling) different minor or major metabolites, either from plants, animal tissues, food, feed or human fluids and cells: gas-chromatography coupled with FID or MS (GC /MS), liquid chromatography coupled with PDA/MS (U(H)PLC/MS) (Sumner, Mendes & Dixon,2003; Dunn & Ellis ,2005), or MALDI-TOF mass spectrometry, Fourier Transform IR (FTIR) Spectroscopy (Ellis, 2003; Fernandez, 2007; Schultz, and Baranska, 2007)), as well capillary electrophoresis or Nuclear magnetic resonance (NMR) (Socaciu 2006, 2008b)

Chemometrics "upgraded" the former biostatistical methods based on multivariate analysis by new methods like Cluster analysis (CA), Principal Component Analysis (PCA) PLS (Partial Least Squares) or Discriminant Function Analysis (DFA). Cluster analysis (CA) is "collecting" the group similar data, "clustered" in associated group members with common properties and differentiate the groups based on their taxonomy, behavour or physiological condition (Lavine, 2006; Bereton, 2007). Data are presented in dendrograms that emphasize the grouping (Fiehn et al., 2000; Roessner et al., 2001). Like CA, Principal Component Analysis (PCA) uses all the metabolite data from a sample to compute an individual metabolic profile that is then compared to all the other profiles. PCA takes the resulting group of data points and rotates it until the maximum variability is obvious. PCA finds the best vectors ('principal components') that give the best sample characterization (Fiehn et al., 2000; Roessner et al., 2001). The data may be represented as two- or three-dimensional plots in which the axes (principal components) which include as much as possible of the total information derived from metabolic variances (loadings). The PLS (Partial Least Squares) method helps in prediction of experimental data against models, prediction of errors from concentration or spectral estimates. In metabolomics, PLS is very useful to validate the Infrared qualitative data against HPLC or GC quantitative data, based on predictions with

pure standards. The Discriminated Function Analysis (DFA) is used to determine which variables discriminate between two or more groups. Computer-generated pair-wise plots of every metabolite in the data set against every other metabolite can be informative (Roessner et al., 2001) especially when hundreds of metabolites are analyzed.

Plant and food metabolomics. The quality of medicinal and aromatic herbs from wild environments shows fluctuations dependent on climate and environmental conditions (Fig.1) , their standardization being a difficult but necessary for checking their authenticity, purity and action (Sumner, 2003; Yadav and Dixit, 2008). The identification of their specific phytochemicals' provide valuable information for their pattern recognition but only if a sufficient high number of samples is analyzed by chemometrics (Maloney, 2004; Bender, 2005) to discriminate among different genotypes and phenotypes or extracts designed for standardized formulas (Liang et al., 2004; Yadav and Dixit, 2008; Giri et al., 2010), according to international legislation.

The evaluation of a herbal or food products by their metabolomic fingerprinting can be accomplished by HPLC with UV(DAD), ELSD, MS detection or GC-MS, HPTLC-densitometry, FT-MIR, NIR, NMR or a combination of these techniques (Fan et al., 2006; Mattoli et al., 2006; Gong et al., 2006; Li et al., 2008; Hashimoto and Kameoka, 2008; Gong et al., 2009; Giri et al., 2010; Contreras, 2010). The Infrared spectrometry, especially the Fourier transform system (FTIR) ia a conviniently used for food fingerprinting, e.g. fruit and juice carbohydrates (1000-1200 cm^{-1}), oil esters (800-1500 cm^{-1}) as well carotenoids (965, 1367 and 1450 cm^{-1}), chlorophylls (1587, 1725 cm^{-1} and wine or tea phenolic derivatives (694-849 cm^{-1}) (Rambla et al., 1998; Patz et al., 2004; Ramasami et al., 2004; Baranska and Schultz, 2006; Fernandez and Agosin, 2007; Schultz et al., 2007;).

Some years ago, based on many measurements on plant and food matrices, we proposed specific metabolomic-metabonomic evaluations (fingerprint and quantification) integrated by a four-steps analysis: UV-VIS spectroscopy (1), Infrared spectrometry (2), GC or HPLC ± FID, PDA or MS detection (3), and chemometrics (4). These analysis assured accurate evaluation of different hydrophilic and lipophilic molecules in plants (medicinal, aromatic plants or crops), food (raw fruits or vegetables, juices, oils) found in wild or cultivated under controlled environment.

III. A systems biology approach: symbiosis of metabolomics & chemometrics

Metabolomics belongs to the growing field of systems biology, but only in symbiosis with chemometrics (Weckwerth, 2003; Van der Greef, 2005). Systems biology has developed in recent years from a technology-driven enterprise to a new strategic tool in Life Sciences (Der Greef, 2007), combining the systems phenotyping with in-depth investigations of biomolecular mechanisms, to understand physiological and/or pathological modifications in a certaion organism (from plants to humans). A prerequisite for having benefits of such a systems approach is a reliable and well-validated bioanalytical platform across complementary measurement modalities, especially transcriptomics, proteomics, and metabolomics, that operates together with a biostatistical/bioinformatics platform.

Considering the necessary symbiosis of metabolomics-based systems biology & chemometrics, we review here some relevant findings of our previous experience in performing analytical methods for metabolomic fingerprinting of fruits and food products, coupled with chemometrics as an integrated, added-value technology for systems biology. Some specific case-studies, relevant for plant and food metabolomics are presented.

MATERIALS AND METHODS

During the last 10 years our department and research center developed series of analytical techniques, from the most available ones (UV-Vis spectrometry) to advanced, precise techniques (HPLC-PDA or LC-MS, GC-FID or GC-MS) and new, rapid FTIR techniques to fingerprint the relevant functional groups (Socaciu et al., 2006, 2007ab, 2008a-e, 2009a-c). We collected data from different fruits specific to hilly regions (seabuckthorn (*Hippophae rhamnoides*), aronia (*Aronia melanocarpa),* bilberry (*Vaccinium myrtilis*), Black currant (*Ribes nigrum*) and characterized their lipophilic or hydrophilic phytochemicals (minor metabolites for pattern recognition) .
Meanwhile we characterized the food products derived, such as oils or juices and followed the study of same phytochemicals to authentify their fingerprint and quality.

We proposed a successive pathway, in four steps, of analytical investigation, using the above-mentioned techniques (Socaciu et al., 2010ab; Pop, 2010, 2011).

First step: UV-Vis fingerprint applied for plant extracts in hydrophilic *vs* lipophilic solvents. The specific absorption maxima gives basic information about the presence, concentration and relative ratios of different molecular groups. The UV-Vis spectroscopy offer a simple, cheap and easy-to-use technique to identify and quantify the main phytochemicals (e.g. carotenoid, anthocyanin, chlorophyll or flavonoid pigments, phenolic derivatives) discriminating between the lipophilic and hydrophilic phytochemicals, in relation to the polarity of the extraction solvent. We used our Perkin Elmer and JASCO spectrometers in the range of 200-700 nm (Socaciu et al., 2005; Vicaş 2010).

Second step: FTIR-ATR (Fourier Transformed Infrared) spectroscopy, which is a rapid and non-destructive investigation, easy to use to fingerprint fresh or dry plant, powders or extracts (Socaciu et al., 2006, 2009bc; Fetea et al., 2008; Chiş et al., 2010; Leopold et al., 2011). The use of attenuated total reflectance (ATR) device evolved rapid FTIR measurements of liquids such as oils and plant extracts, allowing the identification and quantification of valuable plant biomarkers (Schultz and Baranska, 2007).

The third step (validation step) consists in advanced analysis of plant or food extracts (lipophilic or hydrophilic) by GC-FID±MS), HPLC-PDA±MS analysis of extracts, allowing their accurate fingerprint, profiling and quantitation of each class or individual molecules (Dulf et al., 2006, 2008; Ranga et al., 2008; Vicaş et al., 2011).The forth step integrates all previous data by chemometric analysis, for an accurate interpretation and conclusion, including validation of FTIR data by HPLC with (PCA, CA and PLS analysis). Once a good validation of FTIR data is made for a specific plant or food matrix, this method can be repeated easily by routinely for hundreds or samples, providing cheap and reliable results.

Based on these four-steps analysis, specific predictions and evaluations of the quality and authenticity of specific plants or food/feed or animal /human tissues and fluids.

RESULTS AND DISCUSSION

Fig.2 represents the comparative HPLC-PDA fingerprints of lipophilic (A) and hydrophilic (B) extracts from the four fruits investigated. The individual biomarkers (carotenoids and phenolics) are mentioned.C. FTIR fingerprint of the same fruits, showing the regions specific to main phytochemicals.
Fig.3 includes the comparative FTIR spectra (700-3700 cm^{-1}) used to fingerprint seeabuckthorn fruits (SBm) and juice (SBj) (left) or to discriminate between raw and clear juice (right).

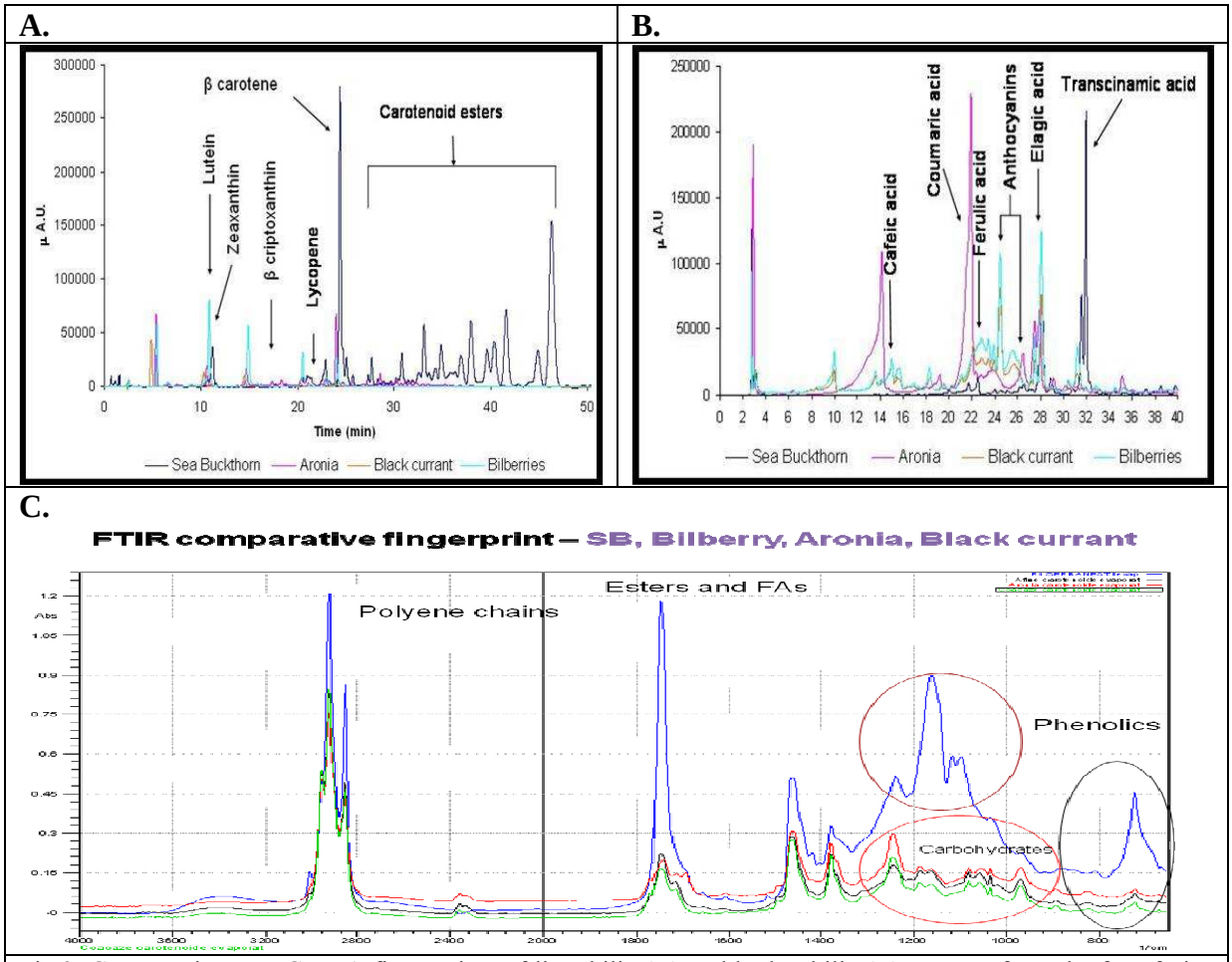

Fig.2. Comparative HPLC-PDA fingerprints of lipophilic **(A)** and hydrophilic **(B)** extracts from the four fruits. The individual biomarkers (carotenoids and phenolics) are mentioned. **C.** FTIR fingerprint of the same fruits, showing the regions specific to main phytochemicals.

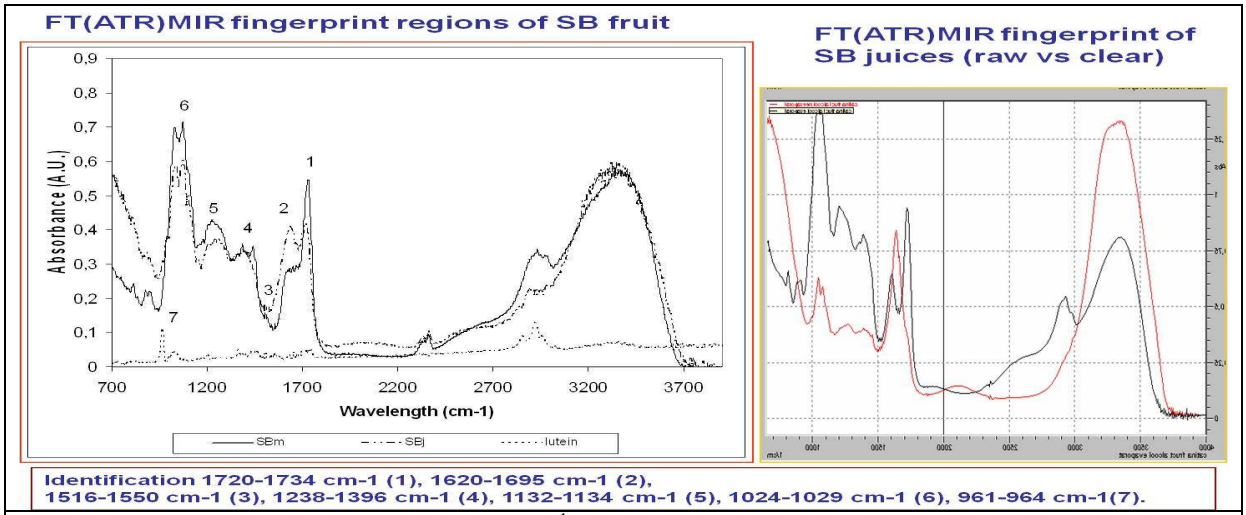

Fig. 3. Comparative FTIR spectra (700-3700 cm^{-1}) used to fingerprint seeabuckthorn fruits (SBm) and juice (SBj) (left) or to discriminate between raw and clear juice (right).

Fig. 4 represents the principal component analysis (PCA) and cluster analysis (CA) based on preliminary HPLC-MS fingerprints of carotenoid and flavonol glycoside from berry vs leaf, for six varieties of seabuckthorn (Pop, 2011).

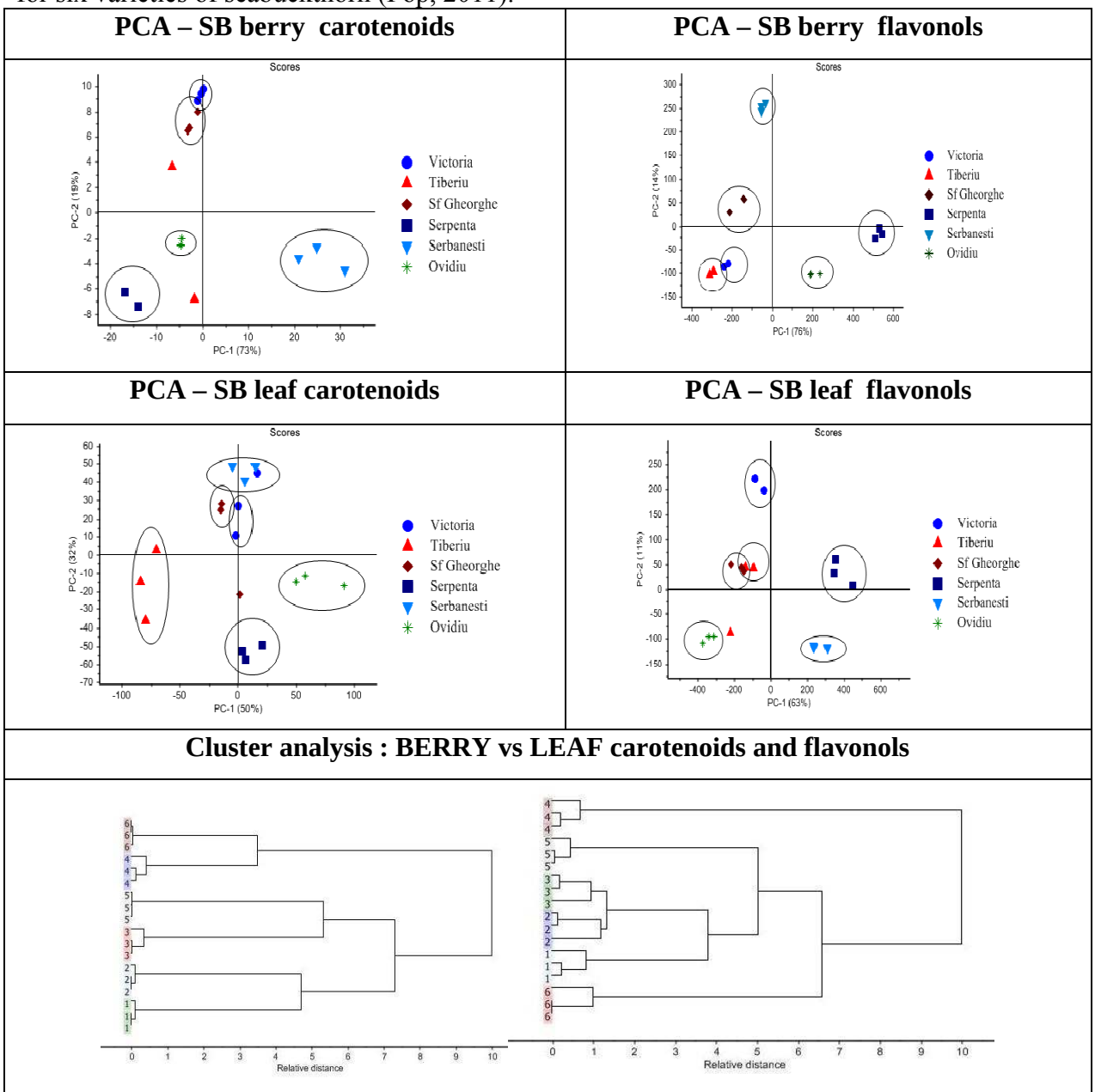

Fig.4. Principal component analysis (PCA) and Cluster analysis (CA) based on HPLC-MS fingerprints of carotenoid and flavonols from berry vs leaf, for six varieties of seabuckthorn.

CONCLUSION

By our specific case-studies we demonstrated here that the determination of biochemical markers for a specific plant or food is just a beginning of a metabolic approach. Only combining the chemical informations given by sophisticated or rapid, simple techniques associated with chemometrics' informations *via* pattern recognition (fingerprinting), data calibration and quantitative measuring, as well clustering of significant groups of samples

based on their principal components are necessary to enrich the metabolic profile and approach the intergrated view of systems biology.

The application areas of Metabolomics are infinite and integrated in the systems biology frame: from plant breeding and crop quality supervision, food quality and safety evaluation, toxicity and nutrition assessments, Pharmaceuticals and drug development, yield improvement in crops and fermentations, genotyping, gene function elucdation and biomarker discovery, technological advances in analytical chemistry, environmental control and adaptations,

Beside characterizing relevant "minor' molecules (phytochemicals) in food/feed products, as authenticity, traceability or quality biomarkers, agrifood (plant & food) metabolomics became a powerful omics' technology with impact in nutrition and health. Meanwhile strong implications in nutrition (nutrigenomics and nutrigenetics) and health (P4 = Predictive-Preventive-Personalized and Participative Medicine).

REFERENCES

1. Baranska, M., H. Schultz (2006). Application of infrared and Raman spectroscopy for analysis of selected medicinal and spice plants. J. Med. Spice Plant. 2:72-80.
2. Bender, D.A. (2005). The promise of metabolomics. J. Sci. Food Agric. 85:7–9.
3. Brereton, R.G. (2007). Applied Chemometrics for Scientists, John Wiley & Sons, Ltd., UK
4. Chis, A. et al. (2010). Application of FTIR spectroscopy for a rapid determination of some hydrolytic enzymes activity on sea buckthorn substrate. Rom. Biotech. Letters.15:5738-5744.
5. Contreras, U. et al. (2010). Screening method for identification of adulterate and fake tequilas by using UV–VIS spectroscopy and chemometrics. Food Res. Int., 43:2356-2362.
6. Dulf, F. et al. (2006). Comparative studies on fatty acid fingerprint from total lipids and phytosterol esters of some edible oils. *Bul. USAMV*. 62:225-231.
7. Dulf, F., et al. (2008). Phytosterol and Fatty acid Profile of Four Edible Oils Processed in Romania. Chem. Listy. 102:s265–s1309
8. Dunn,W.B., D. I. Ellis (2005). Metabolomics:Current analytical platforms and methodologies. Trends in Analytical Chemistry. 24:42-47.
9. Ellis, D., G.G. Harrigan, R. Goodacre (2003). Metabolic fingerprinting with Fourier-transform infrared spectroscopy. In: *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis.* Harrigan G.G., Goodacre R. (eds). Kluwer Academic: Dordrecht.
10. Fan, X. H. et al. (2006). Multiple chromatographic fingerprinting and its application to the quality control of herbal medicines. Analytica Chim. Acta. 555:217-224.
11. Fernandez, K., E. Agosin (2007). Quantitative analysis of red wine tannins using Fourier Transform Mid-Infrared spectrometry. J. Agric. Food Chem. 55:7294–7300.
12. Fetea, F. et al. (2008). FTIR fingerprint of different biological and geographical origin species of Echinacea spp. and Melissa off., Proc. 5th Int.Conf.Plant Metabolomics, Yokohama, Japan
13. Fiehn, O. (2002). Metabolomics - the link between genotypes and phenotypes. Plant Molecular Biology. 48:155-171.
14. Giri, L. et al. (2010). Chromatographic and spectral fingerprinting standardization of traditional medicines: an overview as modern tools. Res. J. Phytochem. 4:234-241.
15. Gong, F. et al. (2006). Variable selection for discriminating herbal medicines with chromatographic fingerprints. Analytica Chim. Acta 572:265-271.
16. Gong, F., Q. Zhang, J. Zhang (2009). Quality Assessment of Herbal Medicine with Chromatographic Fingerprint. J. Drug Addition, Educ. & Eradication, 4: 257-302.
17. Hall, R. et al. (2002). Plant Metabolomics: The Missing Link in Functional Genomics Strategies. Plant Cell. 14:1437-1440.

18. Hashimoto, A. and T. Kameoka (2008). Applications of Infrared Spectroscopy to Biochemical, Food, and Agricultural Processes. App. Spectrosc. Rev. 43:416-451.

19. Lavine, B. and J. Workman (2006). Chemometrics. Anal. Chem.78:4137-4145.

20. Leopold, L.F., H. Diehl, C. Socaciu (2006). HPLC fingerprint of organic acids in fruit juices, Bull. USAMV, 62:288-293.

21. Leopold, L.F. et al. (2011). Quantification of carbohydrates in fruit juices using FTIR spectroscopy and multivariate analysis. Spectroscopy. DOI: 10.3233/SPE-2011-0529

22. Li, S, et al. (2008). Chemical markers for the quality control of herbal medicines: an overview. Chinese Med. 3:1-18.

23. Liang, YZ, P. Xie, K. Chan (2004). Review:Quality control of herbal medicines. J Chromatography. B 812:53-70.

24. Lindon, J.C., J.K. Nicholson, E. Holmes (2007). The Handbook of Metabonomics and Metabolomics. Elsevier, Amsterdam

25. Maloney, V. (2004). Plant metabolomics. BioTeach J. 4:92-99.

26. Mattoli, L. et al. (2006). Metabolomic fingerprinting of plant extracts. J. Mass Spectrometry. 41:1534-1545.

27. Parlog, R. et al. (2009). HPLC-PDA and UV-Vis spectrommetry analysis used to fingerprint Sea buckthorn berries comparatively for leaves and seeds extracts. Bull. USAMV. 66:409-415.

28. Patz, C.D. et al. (2004). Application of FT-MIR spectrometry in wine analysis. Anal Chim Acta 513:81–89.

29. Pop, R., et al. (2010). Comparative qualitative and quantitative evaluation of two metabolic biomarkers (carotenoids and phenolics) of four berries (Seabuckthorn, Aronia, Black Currant and Bilberries) using HPLC and UV-Vis analysis, Proc. Metabolomics, Amsterdam

30. Pop, R. (2011). PhD dissertation (unpublished data).

31. Ramasami, P. et al. (2004). Quantification of sugars in soft drinks and fruit juices by density, refractometry, infrared spectroscopy and statistical methods. South African J. Chem. 57:24–27.

32. Rambla, F.J. et al (1998). Simple partial least squares-attenuated total reflectance Fourier transform infrared spectrometric method for the determination of sugars in fruit juices and soft drinks using aqueous standards. Analyst. 123:277-281.

33. Ranga, F. et al. (2008). HPLC fingerprint of different biological and geographical origin species of *Echinacea spp*. and *Melissa off*. Proc. 5th Int.Conf.Plant Metabolomics. Yokohama, Japan

34. Roessner, U. et al. (2001). Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *The Plant Cell* **13**, 11-29

35. Schultz, H, M. Baranska (2007). Identification and quantification of valuable substances by IR and Raman spectroscopy. Vibrational Spectroscopy 43:13-25.

36. Socaciu, C., F. Ranga and H. Diehl (2005). UV-VIS spectrometry applied for the quality and authenticity evaluation of edible oils from Romania, Bull. USAMV-CN, 61 :295-300

37. Socaciu, C. (2006). Plant and food metabolomics in the post genomic era: concepts, methodologies and applications (review). Bull. USAMV. 62:352-359.

38. Socaciu, C. et al. (2007a). Combined HPLC, UV-Vis, FTIR fingerprints of antioxidant phytochemicals as functional food authenticity, quality and functionality markers. Proc. 3rd Int. Conf. on Polyphenols and Health, Kyoto, Japan .

39. Socaciu, C. et al. (2007b). Quality and authenticity of seabuckthorn oils using successive UV-Vis, FT-IR, NMR spectroscopy and HPLC-, GC- chromatography fingerprints. Proc.Int. Seabuckthorn Assoc. Conf., Quebec

40. Socaciu, C.(ed.) (2008a). Food colorants, Chemical-Functional Properties,CRC Press, USA

41. Socaciu, C. et al. (2008b). Plant and food metabolomics: fingerprint of phytochemicals as authenticity markers for medicinal plant extracts and functional oils. Metabomeeting, Lyon.

42. Socaciu, C. et al. (2008c) Applied Metabolomics in Romania: from plant fingerprint to food quality and authenticity using advanced analytics, 1st Eur. Food Congress, Ljubljana

43. Socaciu, C. (2008d). Agrifood metabolomics - a link between analytical biochemistry and green biotechnologies, Bull. USAMV. 65:410-415.

44. Socaciu, C., F. Ranga, F. Fetea (2008e). Comparative FTIR-ATR/HPLC metabolic profiles and quality assessment of herbal extracts based on some medicinal plants originating from different European countries. Proc.5th Int. Conf. Plant Metabolomics, Yokohama, Japan

45. Socaciu, C. et al. (2009a). Selected biomarkers and methods to fingerprint vegetable functional oils originating from Romania. Proc. Conf. EuroFoodChem XV, Food for Future, Copenhagen, 162-165.

46. Socaciu, C., F. Fetea, F. Ranga (2009b). IR and Raman spectroscopy- advanced and versatile techniques for Agrifood quality and authenticity assessment. Bull. USAMV. 66: 459-465.

47. Socaciu, C. et al. (2009c). Complementary advanced techniques applied for plant and food Authentication. Czech J. Food Sci. 27:S70-S75.

48. Socaciu, C. (2010a). Metabolomic Biomarkers – From Biochemistry To Systems Biology and Personalized Medicine. Proc. Int.Symp. SRBBM Bucharest

49. Socaciu, C. et al. (2010b). Four-steps metabolomic characterization of carotenoid, chlorophyll and anthocyanin pigments as biomarkers of food quality and authenticity. Proc. 6th Int.Congress on Pigments In Food, Budapest

50. Sumner, L.W., P. Mendes, R.A. Dixon (2003). Plant metabolomics: large-scale phytochemistry in the functional genomics era. Phytochemistry. 62:817-836.

51. Van Der Greef J., A. K. Smilde (2005). Symbiosis of chemometrics and metabolomics: past, present, and future J. Chemometrics. 19:376–386.

52. Vicas, S. et al. (2010). The Morphological Features and UV-VIS Analysis of Some Taxonomic Markers of Genus *Viscum,* Bull. USAMV Cluj, Agriculture, 66: 66-69

53. Vicaş, S. et al. (2011). HPLC Fingerprint of Bioactive Compounds and Antioxidant Activities of *Viscum album* from Different Host Trees. Not. Bot. Hortiagrobot. Cluj-Napoca, 39:48-57.

54. Weckwerth, W. (2003). Metabolomics in systems biology. Annu. Rev. Plant. Biol. 54:669-89.

55. Yadav, N.P. , V.K. Dixit (2008). Recent approaches in herbal drug standardization. Int. J. Integr. Biol. 2:195–203.