Develonutri Year 1 report

Project Progress Year 1

The first year has been frenetic with much effort concentrated towards standard analysis and ring testing of the approaches.

WP1 - Chemical standards, inter-lab calibration, validation, ring testing on biological materials.

The first reporting period has seen the vast majority of the underpinning preparative work completed. In WP1, with the responsibility for the chemical standards, inter-lab calibration, validation and ring testing of profiling approaches, progress is as planned. Extensive and detailed studies into (anti)nutrient and micronutrient stability are ongoing and these will inform the more detailed analyses on biological materials (potato, tomato and wheat) with respect to sources of quantitative variability. Following some ranging studies with partner-derived potato, tomato and wheat samples, more extensive ring testing is under way to compare the standard analytical approaches of GC, LC-RI, LC-UV and LC-PDA with the profiling approach GC-ToF-MS and LC-MS. As part of the ring test standard operating procedures have been circulated and these will be rigidly adhered to ensure cross comparability between the technologies. Progress so far with these activities have contributed to the initiation of a large scale laboratory ring-test in which the experimentation has been designed to ascertain the following:

- Can a harmonised extraction procedure can be used to measure all bioactive and anti-nutrients?
- Determine inter-laboratory standardisation and compatibility.
- Determine how robust the quantification procedures used are.
- Elucidate sources of variability. For example at the level of biological material supplied, extraction or the variety of analytical platforms.

With respect to analysis approaches and stability siginficant comfort can be gained from the stability of selected metabolites over time as analysed by GC-MS.

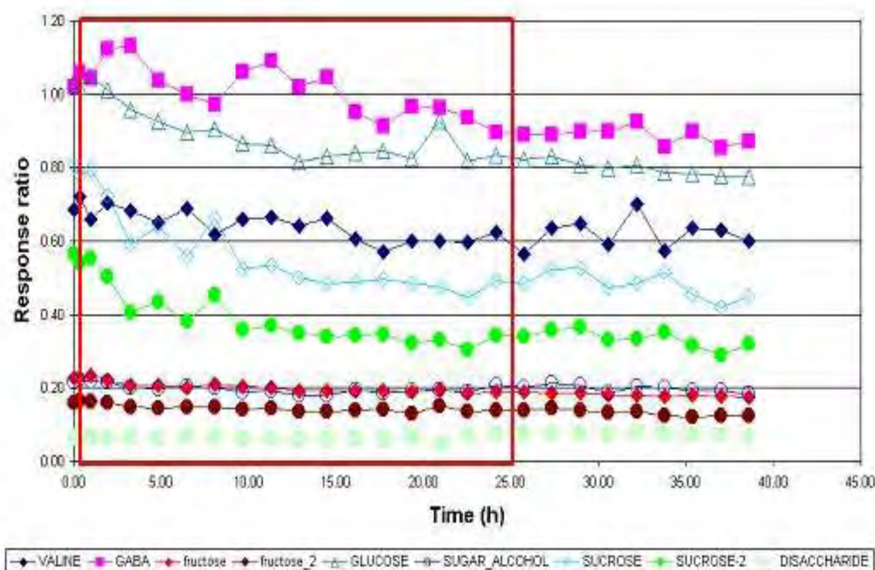

Figure 1: Illustration of selected metabolite GC-MS analysis over time demonstrating the stability of the derivatised samples within the autosampler.

This confidence in metabolite stability meant that we could go ahead and start constructing extensive calibration curves for the nutritional (and putatively bioactive) metabolites via the metabolite approach.

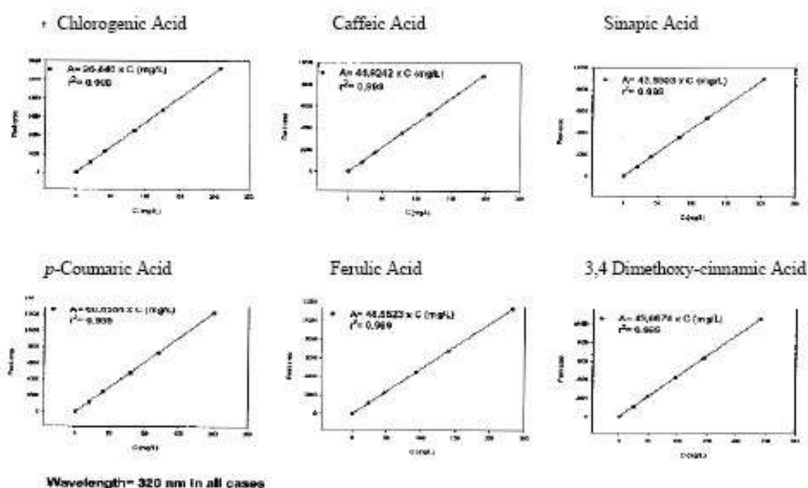

Figure 2: Dose-response curves for selected phytochemicals ascertained using HPLC-PDA.

Clearly the behaviour os standards does not neccesarily reflect how compounds behave in actual food so a series of spiking experiment were undertaken to determine recoveries etc (Table 1 below).

Table 1. (A) The HPLC analyses of carotenoids and (B) water soluble B vitamins. Using an SOP distributed by P5 calibration, limits of detection (LOD) and quantification (LOQ) are reported in the presence of a biological matrix.

(A)

Compound	Range ($\mu g\ mL^{-1}$)	Linearity (R^2)	LOD ($\mu g\ mL^{-1}$)	LOQ ($\mu g\ mL^{-1}$)
Lutein	1.38-6.89	0.9998	0.024	0.079
Zeaxanthine	1.14-5.68	0.9994	0.063	0.211
Canthaxanthine	0.87-4.33	0.9998	0.153	0.510
β-Carotene	0.26-1.30	0.9882	0.167	0.558
Lycopene	np	np	np	np

(B)

Compound	Range ($\mu g\ mL^{-1}$)	Linearity (R^2)	LOD ($\mu g\ mL^{-1}$)	LOQ ($\mu g\ mL^{-1}$)
Ascorbic acid	52.0-259.8	0.9974	0.094	0.310
Nicotinic acid	34.8-174.2	0.9983	0.029	0.100
Pantothenic acid	54.4-271.8	1.0000	0.307	1.020
Pyridoxine	51.3-204.7	0.9941	0.021	0.070
Thiamine	74.6-373.0	0.9999	0.017	0.060
Riboflavin	25.6-101.6	0.9999	0.006	0.020
d-Biotin	58.1-290.5	0.9973	0.038	0.130

These analyses were done with others (listed below) to construct a series of analaytical approach specific standard operating procedures.

- Validation of response linearity vs concentration for a group of standards.
- Validation of response linearity vs amount of plant material extracted for both aqueous and non-polar fractions.
- Validation of response linearity vs volume of aqueous extract derivatised.
- Validation of response linearity vs injection volume.
- Validation of extraction and analysis reproducibility.
- Validation of extraction and analysis reproducibility (multiple injection and/or sample).
- Validation of extract stability whilst on the autosampler tray (where appropriate).
- Estimation of sampling errors, via repeat injection of samples or repeat analyses.

The range of compounds include vitamins, amino acids, carotenoids, tocopherols, phenylpropanoids/flavonoids, glycoalkoids and minerals (micronutrients) using GC-MS, LC-MS and has resulted in optimized methodologies with SOPs for the predominant health-related phytochemicals targeted using traditional methods and more expansive profiling methods using emerging technologies. A characteristic LC-MS profile of a polar tomato extract using a project SOP protocol described above (**P1**) is shown below. The chromatogram is a total ion current response across m/z 80-2000.

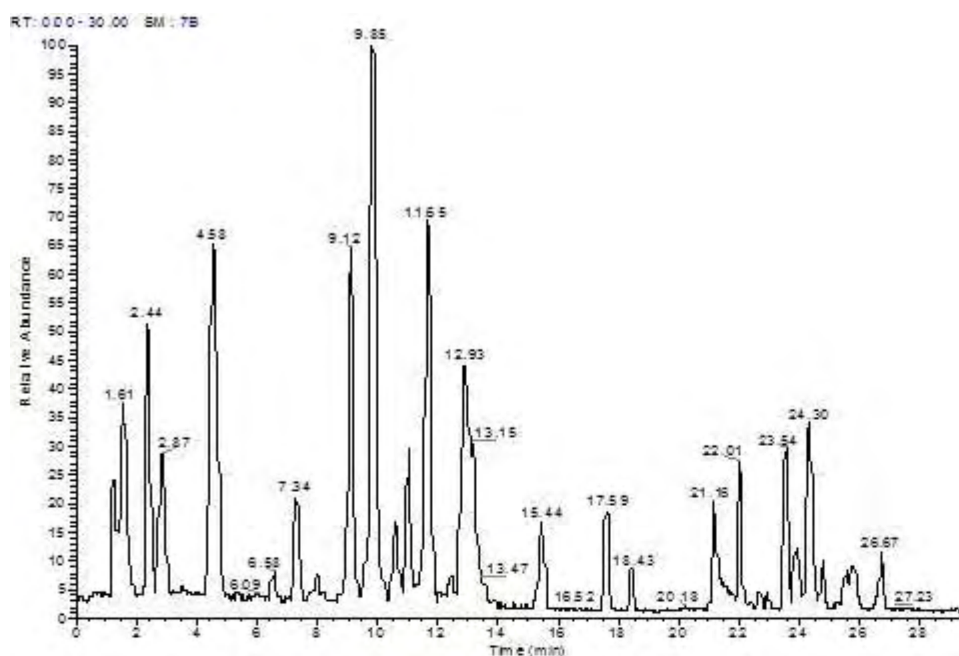

Figure 3: A characteristic LC-MS profile of a polar tomato extract.

A significant level of resources was committed to inter-laboratory ring tests between partners on specific analyses. For example **P2** and **P9** carried out an inter-laboratory comparisons of GC-MS profiling. The results of which suggest that the differences that do occur are mainly associated with the type of hardware used. This has also been found when comparing traditional methodologies where, for example, comparison between Aglient and Waters HPLC systems indicated that the gradient separation for carotenoids or flavonoids could not be reproduced between machines without modification to the gradient (**P1**). This information has had an important bearing on the comprehensive ring-test that is presently underway and highlights the fact that even for what may appear to be a relatively rudimentary part of the analysis, the chromatography, there are sources of variability that would ultimately impinge upon “blind” metabolite/(anti)nutrient quantification.

WP2 - Database Construction and integration of multi dimensional data in an easily accessible format.

In WP2 a database is being constructed that will allow multi format analytical data (UV/Vis, MS, NMR, IR, etc) to be integrated with respect to the crop derived and associated food stuff (anti)nutrient and micronutrients. As part of this interactions with the US National Institute of Science and Technology are ongoing to adapt their Automated Mass Spectral Deconvolution and Identification System (AMDIS) software so that it can be used for LC-MS based approaches. The first iteration of this new software will be relapsed for trial to the partners in project year 2.

All partners have been involved in consultation regarding the construction of the project-derived database both for known (including standards) and unknown chromatogram (or analytical if no chromatography is involved) components. The differentiating criteria will be the analytical platform used. Table 2 illustrates the annotation of the chromatograms used to separate and quantify carotenoids and

tocopherols. Each partner was requested to use the standard metabolomics community system for the annotation of their own chromatograms. For example, laboratory identification, technology, extraction procedure, retention time, spectral properties and identification.

RHU	LC-PDA	EX/PRO	Peak No.	Retention time	Characteristic UV/Vis spectra (λ nm)	λmax III/II	cis peak /λmax II	Carotenoid/Isoprenoid
RHU	LCPDA1	ISOEX1	1	7.84	240	NA	NA	Unk-1
RHU	LCPDA1	ISOEX1	2	8.27	328.0, 406.4, 427.0, 460.8	0.45	0.64	Carotenoid-Unk-1
RHU	LCPDA1	ISOEX1	3	9.85	412.6, 441.5, 470.5	1.07	0	Neoxanthin-1
RHU	LCPDA1	ISOEX1	4	9.90	329.1, 415.0, 440.2, 469.3	0.93	0.06	Neoxanthin-2
RHU	LCPDA1	ISOEX1	5	11.13	326.7, 415.5, 439.1, 468.1	1.05	0.06	Violaxanthin
RHU	LCPDA1	ISOEX1	6	11.70	292.3	NA	NA	α-tocopherol
RHU	LCPDA1	ISOEX1	7	11.92	398.5, 421.0, 448.8	1.1	0	Luteoxanthin
RHU	LCPDA1	ISOEX1	8	12.72	327.0 (406), 427.0, 450.7	0.33	0.3	Carotenoid-Unk-2
RHU	LCPDA1	ISOEX1	9	13.37	326.7, 414.8, 435.5, 463.2	68	0.18	9-cis-Violaxanthin
RHU	LCPDA1	ISOEX1	10	13.57	(278.4), 286.4, (296.2)	NA	NA	cis-Phytoene-1
RHU	LCPDA1	ISOEX1	11	14.22	469.3	NA	NA	Chlorophyll b
RHU	LCPDA1	ISOEX1	12	15.04	332.7(420.1), 445.1, 472.9	68	0.16	Lutein
RHU	LCPDA1	ISOEX1	13	15.95	427.0, 463.2	36	0	Carotenoid-Unk-3
RHU	LCPDA1	ISOEX1	14	16.00	(276.2), 287.6, (298.3)	NA	NA	cis-Phytoene-2
RHU	LCPDA1	ISOEX1	15	16.79	(425.5), 450.0, 479.0	0.42	0	Zeaxanthin
RHU	LCPDA1	ISOEX1	16	17.22	(276.9), 286.4, (298.4)	NA	NA	All-trans-Phytoene
RHU	LCPDA1	ISOEX1	17	17.93	338.6, 417.7, 440.3, 468.1	0.74	0.07	Neoxanthin-3
RHU	LCPDA1	ISOEX1	18	18.66	431.9	NA	NA	Chlorophyll a
RHU	LCPDA1	ISOEX1	19	19.46	331.5, 418.4, 440.3, 468.1	0.60	0.08	Neoxanthin-4
RHU	LCPDA1	ISOEX1	20	19.63	256.8, 333.6, 348.2, 365.8	0.91	0.45	cis-Phytofluene-1
RHU	LCPDA1	ISOEX1	21	19.95	254.5, 331.6, 348.2, 365.8	1.45	0.97	cis-Phytofluene-2
RHU	LCPDA1	ISOEX1	22	20.28	424.2, 442.7, 471.7	0.56	0	Carotenoid-Unk-4
RHU	LCPDA1	ISOEX1	23	20.98	450.0	NA	NA	Unk
RHU	LCPDA1	ISOEX1	24	21.50	278.0, 286.4, 296.8	NA	NA	cis-Phytoene-3
RHU	LCPDA1	ISOEX1	25	21.99	(420.0), 453.6, 477.8	0.43	0	β-cryptoxanthin

Table 2: An example of annotated chromatogram data produced by **P1** from the separation and identification of carotenoids and tocopherols in tomato.

This has been extended to MS properties in the GC & LC-MS systems for the construction of mass spectral libraries for polar and non-polar metabolites of tomato, potato and wheat (durum, bread) has been carried out using freeze-dried material.

Although initial experiments utilized the automated feature of AMDIS library building (for example addition of all extracted component spectra above a certain threshold), it was observed that too many poor quality (visual inspection) spectra were included and as a result this required significant manual intervention (editing or elimination). It was decided, therefore, to apply AMDIS to each data file using default parameters in the first instance and manually add those component mass spectra (extracted, deconvoluted) which appeared to be well defined (generally above a threshold "amount" of 0.001). This was considered to be a satisfactory strategy as the aim was to create "first generation" libraries which will subsequently evolve throughout the project. Related libraries in the National Institute of Science and Technology (NIST) user format have been spawned from the AMDIS libraries and chemical structures have been added in cases where the identity of the component is considered to be "certain" (agreement with at least two other mass spectral databases). This evolution will impact upon or address the following within WP2:

- Assignment of unknowns (some consortium members have greater knowledge of their particular (specialist) crop than others.
- Labelling of artefacts and contaminants.
- Addition of new spectra resulting from ring trials and feedback from other participants.

- Substitution of improved spectra from re-analyses in some cases - for example using a different split ratio in the gas chromatograph to obtain stronger spectra, or fractionating the extract (work in progress) prior to GC/MS analysis where considered beneficial.

The strategy will provide the highest quality mass spectral libraries in a universal format (i.e. independent of instrument manufacturer) of genuine components present in these foodstuffs. If achieved this will represent one of the projects higher aims: a technology independent food compositional mass spectral database.

WP3 - Emerging technologies and their impact as high throughput screening approaches for plant breeding and metabolite and nutrient analysis.

The work package aims to establish how emergent technologies perform as future approaches to high-throughput screening for metabolites and nutrient in the target crops and foodstuffs. These aim to include FT-MS, LC-SPE-NMR, MALDI-ToF-MS, LC-ECD and FT-IR. The approaches will utilise selected biological material on which data has already been acquired via existing and profiling approaches (WP1) to allow for cross comparison.

MALDI-ToF-MS.

Matrix-assisted laser desorption ionisation time-of-flight mass spectrometry (MALDI-TOF/MS) has traditionally been used for the robust accurate determination of peptides derived from trypsinated proteins. A part of **WP3** the utility of MALDI to analyse small molecules has been investigated. One of the fundamental problems of MALDI analysis is the spot to spot variation that results from co-crystallisation of the sample with the matrix. Through the use of a nitrocellulose film impregnated with matrix solution spot-to-spot variation has been alleviate to some extent. The procedure has been applied to polar and non-polar extracts. Polar extracts (methanol) yield signals both in negative and positive mode which constituted a chemical fingerprint and further work is required to characterise the components of this extract. The non-polar extract from tomato in particular gave very strong signals in positive mode and the associated the spectra were dominated by singly charged carotenoids-derived ions. The procedure has been assessed to ascertain if the methodology could be used in a quantitative manner to analyse carotenoids and chemically fingerprint GM genotypes. In conclusion the data show that MADLI-ToF-MS can be used to rapidly and simultaneously profile, identify and semi-quantify plant carotenoids reproducibly, as well as detecting other metabolites in complex biological systems.

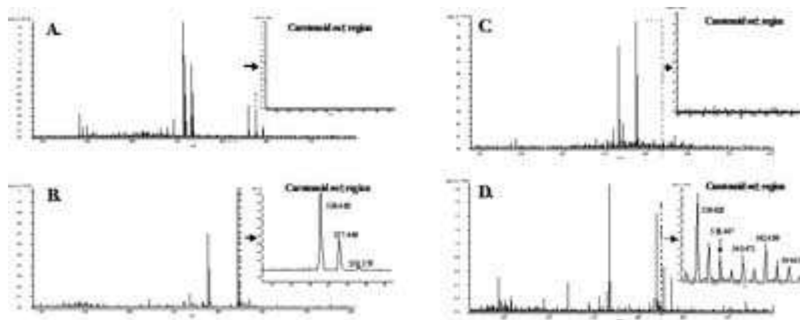

Figure 4: MALDI-ToF-MS profiles of non-polar extracts from; A - Matrix blank with no biological material included in the extraction; B - Non-polar extracts from wild type (Ailsa Craig) ripe tomato fruit, 536.460 *m/z* corresponds to lycopene/β-carotene; C - Non-polar extracts from *Psy*-1 antisense (carotenoid devoid) ripe tomato fruit; D - Non-polar extracts from Tangella ripe tomato fruit.

LC-FT-MS

For evaluation purposes tomato, potato and wheat metabolites were extracted from freeze-dried powders using a solvent partition into an aqueous (polar phase) and a chloroform phase (non polar phase). To assess the reliability of this system in the identification of polar metabolites, we first retrieved the monoisotopic masses of several compounds known to be present in the target tissues. The ChemSpider database was used for this pupose (e.g. ascorbate, citrate, glutamate, tomatine, etc.). These masses were used to extract the putative ion in the various ion chromatograms and determine the mass accuracy. In all cases, the system performed to a high efficiency with a deviation from real masses, in the absence of mass lock, in the range <2 ppm. This highlights the utility of this technology: the ability to obtain measurements of accurate masses (routinely less than 1 ppm) without the need for a mass lock calibration. An example of the total ion chromatograms with the corresponding PDA spectra is shown below the aqueous extracts of tomato (Tic then PDA, potato, bread and durum wheat (negative ion mode).

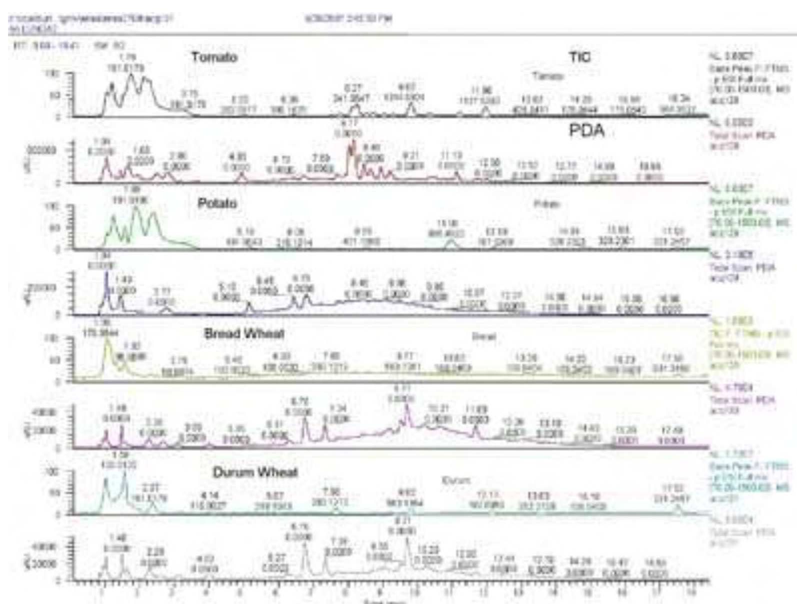

Figure 5: Total ion chromatograms (negative ion mode) from LC-FT-MS with the corresponding PDA spectra for the aqueous extracts of tomato, potato, bread and durum wheat.

Tentative assignment of molecular structures to selected ions was performed by exporting the putative molecular formula and checking its matching score with the monoisotopic masses present in the METLIN database. A direct interface, built into the software, to link the accurate mass with its isotopic mass ratio and a putative molecular structure was not available at the time of our evaluation of this technology but this will be forthcoming.

FT-IR

The application of FT-IR to the rapid analysis and quantification of large and small molecules including metabolites and nutrients is already established for many applications but poorly utilised due to poor experimental design, the overwhelming presence of water, which swamps the signals in the FT-IR spectra, and unfamiliarity with the statistics needed to construct statistically valid regression and quantification curves. **P1** has already published an approach whereby Diffuse Reflectance FTIR (DRIFT) is used to quantify chemical composition in wood[1] and this will largely be duplicated here. Following the analyses by the consortium of the circulated samples these will then have tabulated compositional information. This will be used in conjunction with serial solid sample:solid KBr or NaCL (infrared invisible) dilution regimes for the DRIFT spectra acquisition to establish partial least squares regression lines for each metabolite where possible. Once established these regression curves will be challenged with blind samples and the veracity of the predicted metabolites levels determined against those derived from LC- and GC-based methods

[1] Nuopponen, M., Birch, G.M., Sykes, R., Lee, S. and Stewart, D. (2006) J. Agric.Food Chem. 54, 34-40.: Nuopponen, M., Wikberg, H.I., Birch, G.M., Jããskelãinen, A.-S., Maunu, S.L., Vuorinen, T. and Stewart, D. (2006) J. Applied Polymer Sci. 102, 810-819. Nuopponen, M., Shepherd, T., Birch, G.M., Verrall, S. and Stewart, D (2006) Cell. Chem. Technol. 40, 735-738.

LC-QToF-MS

In the evaluation of this technology, tomato, potato and wheat metabolites were extracted from freeze-dried powders using a solvent partition into an aqueous (polar phase) and a chloroform phase (non polar phase). Aqueous and non-aqeous fractions were analysed using a reversed phase UPLC systems with ESI and atmospheric pressure chemical ionisation (APCI), respectively. An example of the UPLC ToF chromatograms of the tomato aqueous extract (negative ion mode) is shown below.

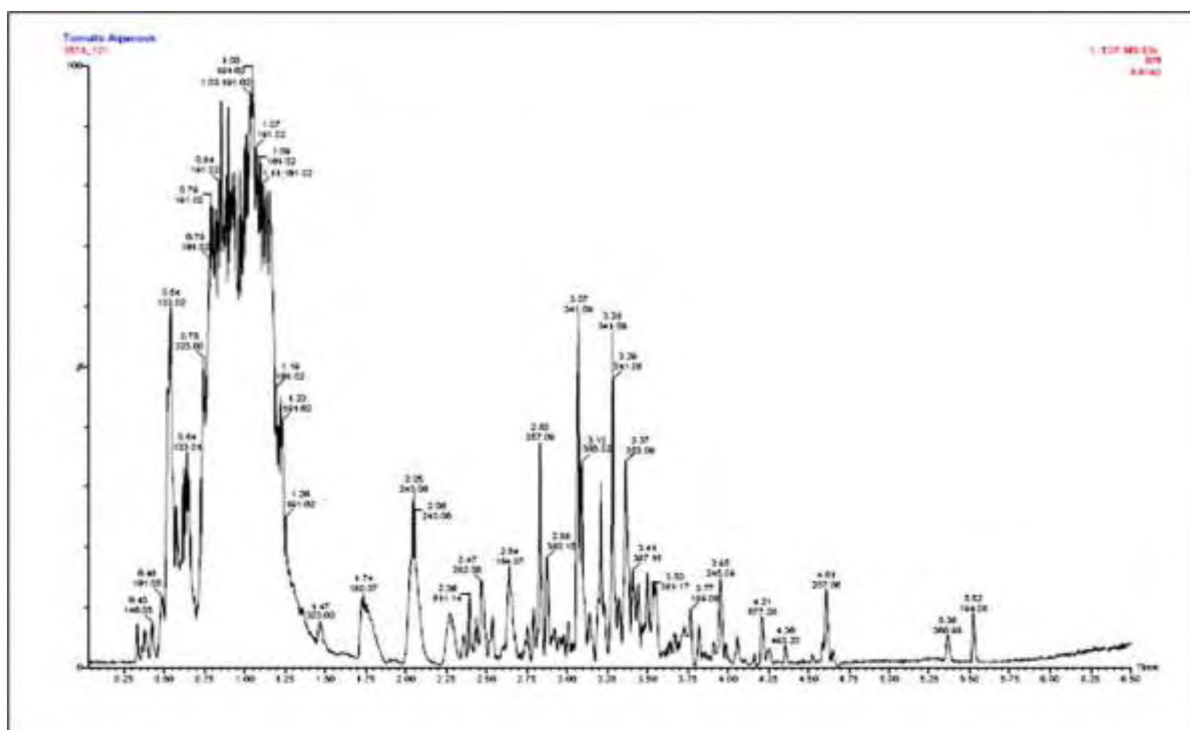

Figure 6: An UPLC-ToF-MS chromatogram of the tomato aqueous extract (negative ion mode).

In order to maximise the identification of the metabolites both PDA (210-500 nm) and MS^n data were recorded. Mass accuracy was improved by the use of “lockspray”, which allowed for automatic verification of the sample mass measurement with the known standard during sample acquisition.

Alignment of the mass chromatograms, normalisation and identification of specific components were performed with the specific software accompanying the instrument (MarkerLynx, MassFRAGMENT) giving a tentative assignment of molecular formula based on isotopic mass ratio of selected peaks. Mass filtering was useful in distinguishing the analyte data from the noise, thus allowing the extraction of a cleaner ion chromatogram. In many cases, following mass filtering and on the basis of elemental composition, the first hit of the molecular formula yielded less than 1 ppm error. The combination of trap and transfer fragmentation of the Synapt High Definition MS (HDMS) System (Waters) yielded very rich MS^n data which will assist in metabolite/mutrient assignment and subsequent machine learned processing and quantification. This approach will be mined in more detail in the following months using samples and extraction protocols that all the partners are using to establish relevance and cross comparability.

WP4 - Provision of biological material and Practical demonstrations.

Central to the progress of the project is the organisation of the experiments for the demonstration activities and provision of biological material for the activities of other WPs which are focussed on the development of standard, state-of-the-art and emergent analytical technologies. The crops for preliminary analysis will be derived from a range of agricultural regimes and locations across Europe and South

Africa. These include GM material for each crop. To capture added value biological material will be gift aided from existing EU projects where possible. Practical demonstration material will be provided via the SMEs. Sampling will occur at all stages of the post-harvest food chain including transportation etc. The materials will be fed into WPs 1-3 across the project duration.

A number of activities have been organized to assure the provision of the biological materials in order to support WP1, 2 and 3 during the second and third year of the project. We decided to select the most appropriate genetic material/farming conditions to capture most of the metabolic variability the European consumer might be exposed to in the open marketplace for the three popular European crops (potato, tomato, wheat [bread and durum]) and their corresponding derived products. For each crop a detailed genotype selection and farming strategy has been defined.

Tomato, cultivars and transgenic lines

Tomato germplasm representing two wild type background varieties (Alisa Craig and Moneymaker), a processing variety (M82), a recently introduce new variety (Kumato) which has had novel food clearance, two cultivars from South Africa (Aroma and Rodate) and two GM varieties that are the prototypes for the GM material formally sold in Europe have been acquired. The GM materials are in the Alisa craig background, and the modification represents down-regulation by either antisense RNA technology or homology induced silencing of a poly-galacturonase (PG) gene. Six plants of each have been cultivated under glasshouse conditions with supplementary lighting in a randomized plot formation. This complete tomato sample set has been distributed to three partners (**P2**, **P3** and **P11**) and each of them is presently growing the material to produce enough fruit for analysis. At the site of **P2** seeds were germinated in December 2007 in the greenhouse and are presently flowering with ripe fruit expected at the end of March and quality fruit expected to persist throughout April.

At the site of **P11** the cultivar set were planted in the field in December and the samples will be ready in between March and April. (No transgenic material was grown in at this site due to the restrictions in GM commercialisation and field sowing). **P3** is currently sowing the tomato seeds in the greenhouse. The variation in sowing time is an inherent consequence of the different geographical regions (South Africa, Southern Italy, UK). Growing conditions will reflect the standard agricultural practices of each location. It is expected that the harvest will each yield ~10kg of fruit per site (the yield of four to six plants) and that these will be freeze-dried (as whole fruit) at each site then sent to the appropriate WP leader for aliquoting and distribution.

Potato cultivars

The selection of the potato cultivars was carried out in coordination with other national and EU projects dealing with this crop where **P1** is/was involved. Six cultivars have been selected to represent the different end uses: Boiling (Desiree and Estima); French Fries (Pentland Dell and Maris Piper); Potato chips/crisps (Lady Rosetta and Saturna). Samples of analogous germplasm (lines) will be collected from other EU (and national) projects where **P1** is also involved and this will allow a

year-on-year and broader environmental assessment on crop derived (anti) nutritional components. Furthermore a new experimental trial will be planted as outlined below in March/April 2008 at the **P1** experimental farm (Figure 13). Samples (chopped, whole tuber) will be freeze-dried at SCRI.

A GM potato line will also be utilized within Develonutri but due to governmental restrictions in GM commercialisation and planting no new material will be planted as part of this project. Within a previous FP6 STREP, NOFORISK: **P1** generated a significant amount of the potato cultivar Desiree that had been down-regulated with respect to the gene Solanosyl glucosyl transferase-1 (SGT-1). The result of this down regulation was a virtual elimination of the glycoalkaloid solanidine but a commensurate increase in chaconine to give total glycoalkaloids levels that were not statistically different from the parent line[2].

Bread wheat cultivar and transgenic lines

Six cultivars have been chosen as representative of the bread wheat genotypes most adapted to European conditions: three Italian wheats, Bilanca, Blasco and Bokaro; two UK wheats, Glasgow and Mascot and one Swiss wheat, Zinal. Furthermore two GM lines and the corresponding controls have been selected to address this issue. These are a Frisal wheat carrying the selectable marker bar and expressing two proteins with putative antifungal effects (chitinase and ß-1, 3-glucanase) and a Bobwhite wheat carrying the selectable marker phosphomannose isomerase and expressing either the allele *Pm3a* of a resistance gene against powdery mildew or the allele *Pm3b*. The associated control lines are also being grown alongside the GM material and will be harvested accordingly. Growing conditions will be according to the standard agricultural practices in each location.

The (non-GM) wheat cultivars were sown during autumn 2007 in three locations representative of different European wheat growing areas: **P1** - Scotland; **P4** - Switzerland and **P5** - Southern Italy. From each location about 1kg of seeds will be collected, milling will be centralised (**P5**) whereupon it will be sent to the appropriate WP leader for distribution. At the experimental farm of **P1** the wheat trials have been constructed to take two levels of fertilisation into consideration.

Since the GM wheat lines are both derived from spring-sown cultivars time the experimental field will be sown in March 2008 in Swiss where an authorized location for GM trials was established in the previous year by **P4**. Milling of GM seeds and controls will be carried out by **P4**. Both non-GM and GM genotypes will be harvested in summer 2008.

Durum wheat cultivars

Since durum wheat is traditionally grown in Southern Europe and mainly in Italy, the cultivars chosen are among the most adapted to the Mediterranean environment and their site selection is based on the data from last 15 years Italian national durum wheat national trials. The following cultivars have been selected: Creso, Simeto, Svevo, Saragolla, Sfinge and PR22D89. These genotypes are a good representation of the most popular, high yielding/high quality durum wheat cultivars

available. The six cultivars were sown in three locations across Italy (from South to North) selected from amongst the locations in the national field trials list.

One location is the experimental farm of P5, where the same cultivars will be also grown under an organic farming system (side–by-side with the standard farming system). Growing conditions will be according to the standard agricultural practices in each location. From each location ~1kg of seed/cultivar will be harvested, milled centrally at P5 and subsequently distributed to the consortium partners. The samples will be harvested in early summer 2008.

[2] McCue, K.F., et al 2007. Potato glycosterol rhamnosyltransferase: the terminal step in triose side-chain biosynthesis. **Phytochemistry 68**, 327-334.

WP5 - Management and outreach.

In WP5 the organisational issues including IP, outreach, knowledge transfer etc are being dealt with As an example of the knowledge transfer aspect the project has already had been highlighted and discussed on several UK national television and radio interviews. A project website is being built at www.develonutri.info. In addition, a broad brush article on the project was published in the journal of the European Parliament.