Agricultural Innovation Program
Research Project Final Report
Contribution Agreement - Vote 10 Funding

Project Title:	Development of a Designer Soybean Testing Methodology Activity 2: Compound/Constituent Identification in Raw Soybeans as well as in Finished Products
Start Date (yyyy-mm-dd):	2012-04-01
Expected End Date (yyyy-mm-dd):	2013-03-31
Actual End Date (yyyy-mm-dd):	2013-03-31
Principal Investigator (PI):	Sevita International – Jim McCullagh

Short Executive Summary of report:

ECODA's major soybean industry partner, Sevita International, exports a wide range of soybean varieties to Japan for use by Japanese processors to make soymilk and tofu. Each Japanese customer has a different formulation, process and style of product and, therefore, different soybean varieties work better for some customers than others.
In order to identify germplasm suitable to produce preferred soymilk and tofu varieties for international markets, a study of the metabolome of the soybean was initiated using a state of the art ultra-high performance liquid chromatography with time-of-flight mass spectrometry (UPLC MS QTOF) analytical facility.
Raw soybean seeds were analyzed to identify novel compounds within the seed and then end products (soymilk and tofu) were analyzed to determine which compounds in the raw soybean seeds persisted during the production process and could be identified in the end products.
This activity has resulted in methods developed for the first time to provide a detailed, non-targeted analysis of extractable, small molecular weight compounds in soybeans, many of which are primarily involved in taste and odour perception, and also several hundred compounds that were uniquely annotated (using mass to charge ratio (m/z) and retention time (rt) data). Most of the major compounds (>200) have been identified chemically with certainty using standards and internet references.
Six biomarkers for soymilk and three biomarkers for tofu have been discovered. Compounds found in both the raw soybean seeds and the end products can now be used for the screening of Sevita International's germplasm to determine if there are any particular varieties that can be identified for end user testing and future sensory evaluation studies.

A. Research Progress and Accomplishments (to date in relation to expected milestones and deliverables / outputs)

- Include brief summary of:
  - Introduction, literature review, objectives, milestones and deliverables / outputs.
  - Approach / methodology (summary by objectives).
- Include results and discussion (overview by objectives and milestones), next steps and references.

Introduction

The advent of metabolomics, the last of the "omics" technologies following genomics, transcriptomics and proteomics, brings a powerful new technology to the study of soybean varieties and food products based on real constituents. Metabolomics is the systematic study of the unique chemical fingerprints that specific cellular processes leave behind. Metabolomics begins by gathering information on all the small molecule constituents of the sample in an "untargeted" approach, that includes all metabolites present rather than a few "targeted" compounds analyzed in traditional approaches. Metabolite profiling and fingerprint analysis are now being used to identify potential biomarkers capable of distinguishing different species, varieties and commercial products with the aim of establishing quality control code protocols based on biochemical phenotype (Lee et al. 2009). Improvements in technology have made this possible through greatly advanced metabolite separation called "ultra high pressure liquid chromatography", which can resolve over a thousand molecules in less than 10 minutes. Not only can molecules be "tagged" by traditional Mass Spectrometry (MS) techniques, but modern metabolomic facilities use a high resolution MS analysis, isotope algorithms and fragmentation patterns to identify the

metabolites without need for costly and often unavailable pure standards. Next, bioinformatic methods are applied to the identified metabolome to provide statistical comparison or differentiation of soybean materials. Comparison is usually achieved by multivariate analysis such as principal component analysis. An algorithm is used to determine the similarity of samples based on all analysed metabolites. This similarity between samples is then plotted using advanced analytical software (MassLynx and MarkerLynx, Waters Inc. MA, USA). Conveniently, this software allows for a 2D- or 3D-graphical representation of principal component analysis to assist in identifying similarities between samples. Alternatively, the Scores plot (S plot) can be used in a two way comparison of two soybean materials, to determine on a statistical basis, the few biomarkers that can differentiate each material.

Applications of metabolomics include profiling of elites for a large number of metabolites, identifying germplasm comparable to known "elite" germplasm using metabolomics for comparisons and identification of metabolomic "biomarkers" for key characteristics. The ability to identify a large number of compounds in the metabolite profile facilitates the identification of unique components and metabolic features of a cultivar. The results are immediate and comprehensive. Previous targeted approaches, based on a limited number of metabolites, may miss key features. While other "omic" technologies can assess the similarity or difference between germplasm types based on genetics, metabolomics has a distinct advantage in the food area by focusing on the small molecular weight compounds that contribute most directly to taste and odour in soybeans. Here, the Principal Component Analysis (PCA) method will be used to compare known "elite" Canadian and international commercial varieties with Sevita International's extensive germplasm collection. This facilitates the identification of closely matching varieties, based on the metabolome, and distinguishes varieties that are far away from the elite. Comparison of known varieties with useful characteristics, such as good tofu characteristics, with other non-performing varieties for these traits using the S plot, will lead to identification of the biomarkers for good performance and also identify biomarkers that may reduce performance. Once biomarkers are identified, other germplasm can be assessed by the method to identify potential high performance varieties that were previously unknown.

Brenda Simmons 13-5-30 7:20 PM
Comment [1]: Are the Canadian and international varities?

Brief literature review

Previous research on the soybean metabolome has mainly taken a targeted approach to different groups of compounds. For example, Berhow et al. (2002) have identified methods for the analysis of major soybean saponins in processed products. Kaneko et al. (2011) have identified key volatile odour compounds in soymilk. Phenolics, such as anthocyanins and flavonoids, have been extensively studied by Kovinich et al. (2011, 2010) and others.

Objectives

- 2.1. Compound/constituent identification in raw soybean seeds: perform non-targeted compositional testing on soybean varieties that are either preferred or not preferred by soy-food manufacturers.
- 2.2. Compound/constituent identification in finished product: perform non-targeted, compositional testing on soy-food products made with known soybean varieties.

Deliverables

- A summary of findings.

Method

1. 80% methanol extraction
    1. Grind 2g of soybean seed in a Wiley mill (Arthur H. Thomas Co.) through a 1mm mesh (size 20).
    2. Weigh accurately 1.0g of ground material and extract with 10ml of 80% methanol (bulk, Fisher Scientific) by shaking at 200rpm for 60min.
    3. Centrifuge at 4000rpm, 10min, 25°C (Eppendorf, 5810R) and decant the supernatant.
    4. Re-extract residuals with 5ml of methanol by shaking at 200rpm for 60min.
    5. Repeat step 1.3 and then combine the supernatants.
    6. Evaporated the supernatant to dryness by blowing air in a fume hood.
2. Hexane wash and preparation for UPLC-Q-TOF analysis
    1. Re-solubilize dried extract in 1ml of 80% methanol (LCMS grade, Fisher Scientific) by sonication for 5min in a glass vial.
    2. Add 1ml hexane (optima grade, Fisher Scientific) to the vial, hand shake the vial three to five times, wait 15min for separation.

3. Transfer the hexane layer to another glass vial.
4. Repeat steps 2.2 and 2.3 two times.
5. Adjust the final volume of the 80% methanol layer to 5ml with 80% methanol.
6. Filter 1ml of extract into a HPLC vial by PTFE syringe filters (0.22 mm, Whatman).
7. Prepared sample are stored at –20 °C until analysis.

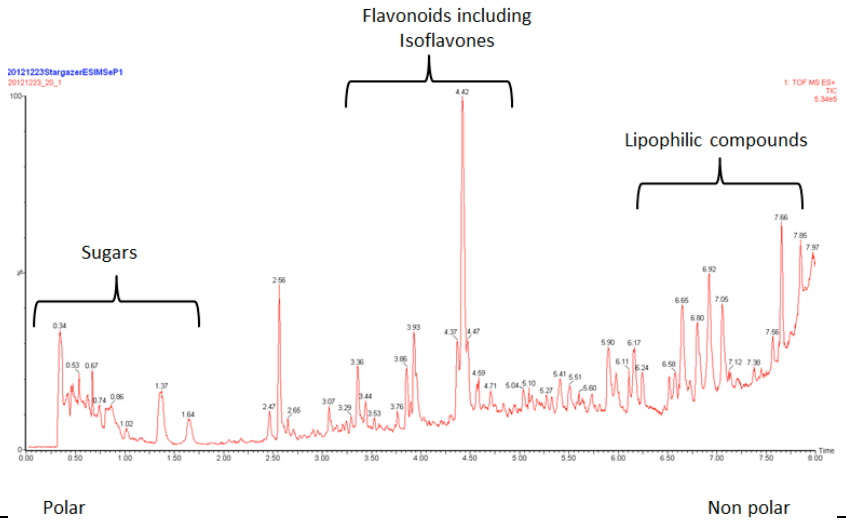
3. UPLC-QTOF analysis

- Optimized UPLC conditions: Acquity CSH C18 1.7um 2.1x100mm column (part #186005297; lot #0113320401) connected with a VanGuard Pre-column 2.1x5mm.
- Mobile phase, A, water+0.1% formic acid, B-acetonitrile+0.1% formic acid (Optima LC-MS, Fisher Scientific).
- Flow rate: 0.5 ml/min.
- Column temperature: 50°C, sample temperature: 25°C.
- Mobile phase composition: 0-1 min 5% A isocratic, 1-6 min linear gradient 5-50% B, 6-8 min 50-95%B, 8.01-10 min 5% A isocratic (total run time 10 min).
- Optimized sample injection conditions: 1ul injection, weak wash 600ul (10% acetonitrile+90% water), strong wash 200ul (90% acetonitrile+10% water).
- Optimized QTOF analysis conditions: MassLynx software, MSe ESI+ mode, lock mass Leucine Enkephalin ^{12}C 556.2615, source temperature: 120°C, desolvation temperature: 400°C, cone gas (N2) flow: 50 l/hr, desolvation gas (N2) flow: 1195 l/hr. MSe conditions, mass range 100-1500 Da, F1 CE, 6V, F2 CER 10-30V, cone voltage: 20V, scan time: 1 sec. calibration, 50-1000 Da sodium formate.
- Optimized statistical analysis conditions: Principal Component Analysis (PCA) and discriminate analysis (OPLS-DA) were performed by MarkerLynx. Pareto scaling was performed after grouping all the samples.

Results and Discussion - Objective 1: Compound / constituent identification in raw soybean seeds

A standard operating procedure was developed for extraction of raw seeds from 76 soybean varieties provided by Sevita International. Analytical methodology for soybean seed metabolome was developed and optimized.
Seeds of raw soybean varieties, submitted by Sevita International, were subjected to metabolome analysis by the validated UPLC-QTOF. Figure 1 shows a representative chromatogram with the three major compound groups identified.

Figure 1: Chromatographic separation of major components of the raw soybean seed



As an example of the utility, Figure 2 shows a representative chromatogram of a soybean variety with 42 annotated markers out of 300 compounds that were identified from 5000 detected metabolites in raw materials and finished soybean products by automated online search in Chemispiderman, Metlin and Plant Metabolite Network. The method could be transferred to HPLC-MS for routine, rapid and inexpensive and more accessible identification of selected group of markers.

Figure 2: Chromatographic separation of soybean 80% methanol extract

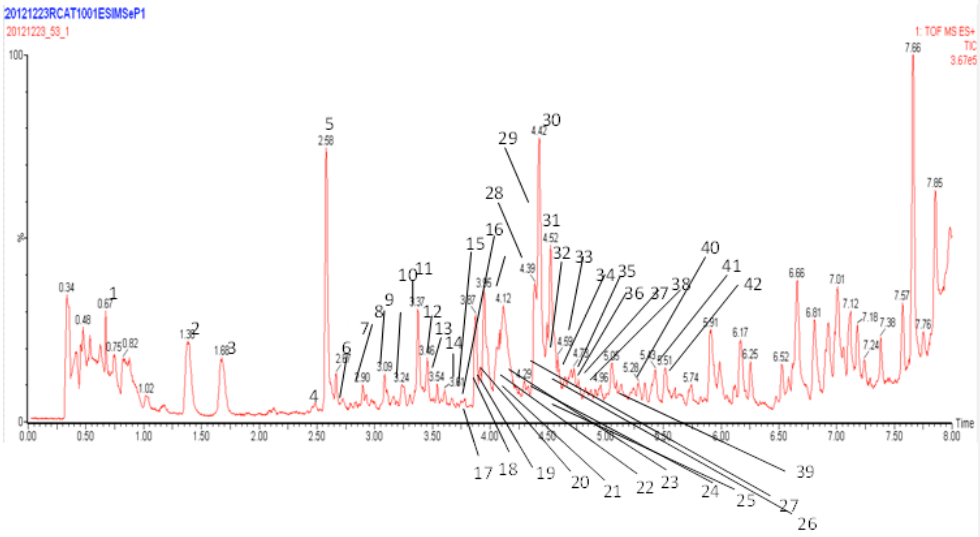
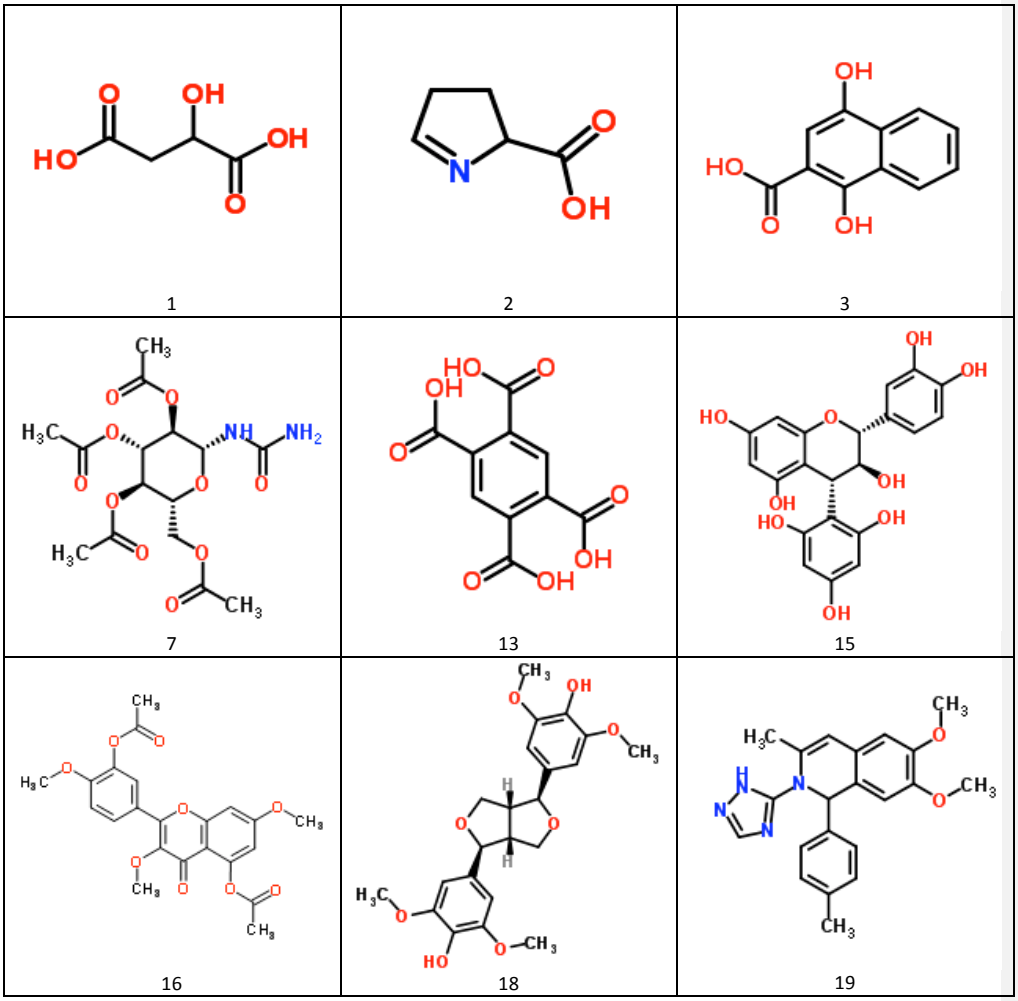


Figure 3 shows the structures of selected biomarkers identification in raw seed of a soybean variety. Only 62 compounds were previously reported in literature. This is the first report of its kind for the soy metabolome. This is a major advance that increases the known metabolome substantially. It provides much greater coverage of compounds important in quality such as taste, odour, resistance to pests, industrial processing and - most importantly - medicinal and nutraceutical benefits such as antioxidants, anticancer and anti-Alzheimer's activity.

Figure 3: Representative example of biomarkers identified in raw soybean seed by UPLC-QTOF



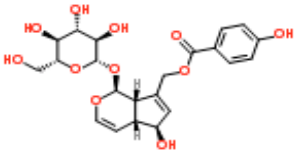
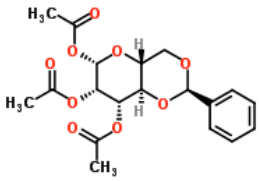
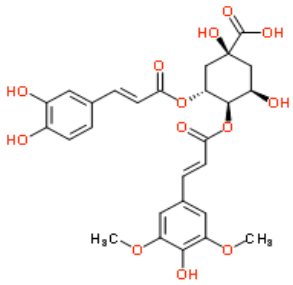
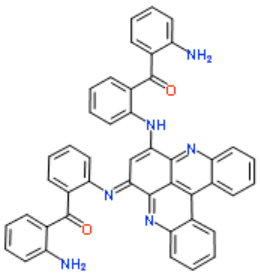
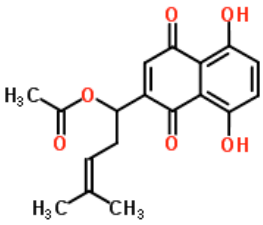
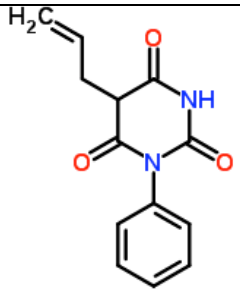
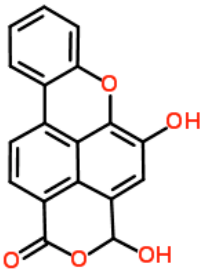
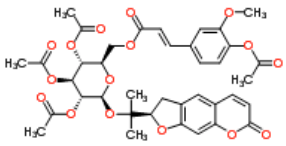
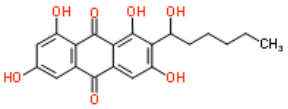
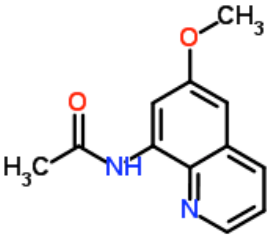
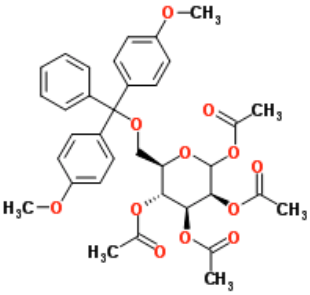
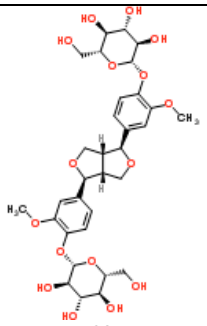
 22	 23	 24
 25	 28	 29
 30	 33	 34
 35	 36	 38

Table 1. Identification of compounds in soybean samples by QTOF (markers in bold have been previously reported to be present in soybean)

Rt (min)	m/z	Elemental composition (PPM, iFIT)	Common/chemical name (date)	Class
0.369	112.9433	C4H2O4	2-Butynedioic acid	
0.44	**173.0546**	**C6H3O6**	**cis-Aconitate**	
0.443	154.0127	C6H9N3O2	Histidine	Amino acid
0.448	127.0072	C5H8N2O2	Dihydrothymine	
0.459	**103.9959**	**C3H7NO3**	**Serine**	Amino acid
0.459	284.0704	C19H11NO2	3-Phenylbenzo[g]isoquinoline-5,10-dione	quinoline
0.46	174.0416	C7H4N5O	N-(4,5-Dicyano-1H-pyrazol-3-yl)acetamide	
0.46	112.9929	C5H6O3	2-oxopent-4-enoic acid	
0.461	366.1428	C15H20N5O6	2-[(2-Amino-6-oxo-3,6-dihydro-9H-purin-9-yl)methoxy]-1,3-propanediyl dipropionate	
0.463	118.0094	C4H9NO3	homoserine	Amino acid
0.463	**189.0395**	**C7H9O6**	**3-Dehydroquinate**	
0.469	225.0062	C10H10O6	Chorismic acid	Amino acid
0.469	127.9914	C5H7NO2	3-Hydroxypyrroline-5-carboxylate	
0.471	**146.0008**	**C5H7NO4**	**L-threo-3-methyl-aspartate**	
0.477	203.0562	C9H7N4O2	1-(3-Phenyl-1,2,4-oxadiazol-5-yl)urea	
0.478	517.1134	C28H21O10	Cefalochromin	
0.486	**181.0234**	**C9H10O4**	**3-Methoxy-4-hydroxyphenylglycolaldehyde**	
0.49	374.1099	C15H21NO10	2,3,4,6-Tetra-O-acetyl-N-formyl-β-D-glucopyranosylamine	
0.49	131.9866	C4H7NO4	Aspartic acid	
0.502	458.1321	C20H21N5O8	2',3',5'-Tri-O-acetyl-8-(2-furyl)adenosine	
0.52	195.0019	C6H11O7	gluconate	
0.53	267.022	C5H7N4O9	4-Methoxy-1,1,3,3-tetranitrobutane	
0.539	**158.0001**	**C7H11NO4**	**N-acetyl-2-glutamate 5-semialdehyde**	Amino acid
0.55	**149.0004**	C5H10O5	**D-Lyxose**	Sugar
0.564	799.1686	C37H36O20	2-(3,4-Diacetoxyphenyl)-4-oxo-3-[(2,3,4,6-tetra-O-acetyl-beta-D-glucopyranosyl)oxy]-4H-chromene-5,7-diyl diacetate	
0.623	**134.9872**	**C5H10O5**	**D-Arabinose**	Sugar
0.625	**158.9847**	**C6H6O5**	**2-Oxoadipate**	
0.641	**208.9807**	**C6H8O8**	**Glucarate**	
0.671	**132.9706**	**C4H6O5**	**Malic acid**	
0.68	**114.9618**	**C4H2O4**	**Maleate**	
0.731	341.0511	C15H9N4O6	5-(2,4-Dioxo-1,3,4,5-tetrahydro-2H-chromeno[2,3-d]pyrimidin-5-yl)-2,4,6(1H,3H,5H)-pyrimidinetrione	
0.739	**172.0126**	**C7H11NO4**	**N-Acetylglutamate semialdehyde**	
0.743	**130.0094**	**C5H9NO3**	**2-oxo-5-aminovalerate**	

0.801	**190.9714**	**C6H9O9**	**2,3-Dioxogulonate**	
1.009	**164.0257**	**C6H7N5O**	**1-methylguanine**	
1.025	**110.9598**	**C5H3O3**	**2-Furoate**	
1.17	**186.9781**	**C7H8O6**	**homo-cis-aconitate**	
1.177	**180.0181**	**C9H11NO3**	**Tyrosine**	
1.251	**204.9858**	**C7H10O7**	**Homocitrate**	
1.33	**315.0701**	**C13H16O9**	**3-Hydroxy-2-methoxyphenyl beta-D-glucopyranosiduronic acid**	Isoflavone
1.37	**657.1066**	**C30H25O17**	**(1S)-1,5-Anhydro-1-(5,7-dihydroxy-2-methyl-4-oxo-4H-chromen-6-yl)-2,3-bis-O-(3,4,5-trihydroxybenzoyl)-D-glucitol**	Isoflavone
1.447	**130.9936**	**C5H7O4**	**(S)-2-acetolactate**	
1.598	**116.0075**	**C4H7NO3**	**Aspartyl semialdehyde**	Amino acid
1.627	218.0525	C5H8N5O5	1-(3,5-Dinitro-1,3,5-triazinan-1-yl)ethanone	
1.637	**225.0249**	**C7H5N4O5**	**3,5-Dinitrobenzohydrazide**	
1.674	407.1139	C23H20O7	10,11-Dimethoxy-3,3-dimethyl-3H-pyrano[2',3':7,8]chromeno[3,2-b][1,5]benzodioxepin-7(14H)-one	
1.873	**184.0135**	**C7H7NO5**	**2-Amino-3-carboxymuconate semialdehyde**	**Amino acid**
2.064	151.9665	**C3H6NO4S**	**3-sulfinoalanine**	
2.152	259.0764	C18H11O2	**2,5-Diphenyl-1,4-benzoquinone**	
2.2	367.1023	C17H19O9	Methyl (1S,3R,4R,5R)-3-{[(2E)-3-(3,4-dihydroxyphenyl)-2-propenoyl]oxy}-1,4,5-trihydroxycyclohexanecarboxylate	
2.315	**174.9059**		**Dihydrogen diphosphate**	
2.339	245.056	C12H9N2O4	5-(4-Methoxybenzylidene)-2,4,6(1H,3H,5H)-pyrimidinetrione	
2.377	161.0355	C8H5N2O2	**5-Nitro-1H-indole**	
2.435	403.1043	C20H19O9	**5-Hydroxy-2-(4-hydroxy-3-methoxyphenyl)-3,6,7,8-tetramethoxy-4H-chromen-4-one**	Flavonoid
2.447	158.0366	C3H6NO4S	**Indoleacetaldehyde**	
2.458	215.0332	C12H12N2O2	8-Acetamido-6-methoxyquinoline	Quinoline
2.554	147.0018		2-Dehydro-3-deoxy-D-arabinoate	
2.556	625.1168	C30H25O15	**2-(3,4-Dihydroxyphenyl)-5,7-dihydroxy-4-oxo-4H-chromen-3-yl 6-O-[(2E)-3-(3,4-dihydroxyphenyl)-2-propenoyl]-beta-D-glucopyranoside**	Flavonoid
2.655	422.0985	C21H16N3O7	**rigidin D**	Alkaloid
2.705	245.0591	C17H10O2	(3Z)-3-(3-Phenyl-2-propyn-1-ylidene)-2-benzofuran-1(3H)-one	
2.727	389.1217	C20H21O8	2-(Hydroxymethyl)phenyl 6-O-benzoyl-beta-D-glucopyranoside	Flavonoid
2.745	264.0406	C14H6N3O3	11-Hydroxy-3H-benzo[h]pyrido[4,3,2-de]cinnoline-3,7(2H)-dione	
2.759	259.0201	C8H7N2O8	3,6-Dinitro-4-cyclohexene-1,2-dicarboxylic acid	
2.763	**144.001**	**C5H7NO4**	**2-oxoglutaramate**	

2.778	**431.0985**	**C21H19O10**	**(1S)-1,5-Anhydro-1-[5,7-dihydroxy-2-(4-hydroxyphenyl)-4-oxo-4H-chromen-6-yl]-D-glucitol**	Flavonoid
2.784	409.1116	C19H21O10	Benzyl 2,3,4-tri-O-acetyl-alpha-D-glucopyranuronate	
2.84	389.12	C15H21N2O10	2,3,4,6-Tetra-O-acetyl-N-carbamoyl-beta-D-glucopyranosylamine	
2.904	**223.012**	**C11H11O5**	**sinapate**	
2.917	**157.0067**	**C4H6N4O3**	**Allantoin**	Alkaloid
2.927	422.098	C21H16N3O7	6-(4-Hydroxy-3-methoxybenzoyl)-5-(4-hydroxy-3-methoxyphenyl)-1H-pyrrolo[2,3-d]pyrimidine-2,4(3H,7H)-dione	
3.002	120.987	C7H5O2	Benzoate	
3.019	221.0342	C13H5N2O2	6,7-Dicyano-2-naphthoic acid	
3.024	331.0087	C15H7O9	5-Methoxy-4,10-dioxo-4H,10H-pyrano[2,3-f]chromene-2,8-dicarboxylic acid	
3.051	409.0939	C22H17O8	3-(2,3-Dihydro-1,4-benzodioxin-6-yl)-2-methyl-4-oxo-4H-chromene-5,7-diyl diacetate	Flavonoid
3.054	431.0757	C24H15O8	(6aR,6bS,13bR,13cR)-13,14-Dimethoxy-6a,6b,13b,13c-tetrahydro-6H,10H-chromeno[6'',7'':4',5']furo[3',2':3,4]cyclobuta[1,2-c]furo[3,2-g]chromene-6,10-dione	
3.103	215.0468	C12H12N2O2	8-Acetamido-6-methoxyquinoline	
3.126	179.0111	C9H8O4	2,4-dihydroxycinnamate	
3.151	**197.0693**	**C9H9O5**	**3,4-Dihydroxyphenyllactate**	
3.174	551.1206	C28H23O12	4-Acetyl-2-(4,5-dihydroxy-2-methyl-9,10-dioxo-9,10-dihydro-1-anthracenyl)-3,5-dihydroxyphenyl beta-D-xylopyranoside	
3.216	**133.0221**	**C5H9O4**	**2,3-Dihydroxyisovalerate**	
3.277	757.2017	C36H38O18	5-(3-Acetyl-2,6-dihydroxy-4-methoxyphenyl)-8-hydroxy-6-methyl-9,10-dioxo-9,10-dihydro-1-anthracenyl 6-O-beta-D-glucopyranosyl-beta-D-glucopyranoside	Flavonoid
3.277	775.2472	C37H43O18	Dimethyl (2aR,4R,4aR,5R,7aS,8S,10R,10aS,10bR)-5,10-diacetoxy-4-methyl-8-{[(2E)-2-methyl-2-butenoyl]oxy}-4-[(1S,8R,9S,11R)-9-methyl-3,7-dioxo-2,4,10-trioxatricyclo[6.3.1.0~9,11~]dodec-11-yl]-3-oxooctah	
3.281	**387.1067**	**C20H19O8**	**5-Hydroxy-6,7-dimethoxy-2-(3,4,5-trimethoxyphenyl)-4H-chromen-4-one**	Flavonoid
3.301	**172.0508**	**C9H6N3O**	**2-Quinoxalinecarboxamide**	
3.334	**179.9818**	**C8H7NO4**	**(4-Nitrophenyl)acetic acid**	
3.354	**252.998**	**C10H5O8**	**1,2,4,5-Benzenetetracarboxylic acid**	
3.363	**113.019**	**C6H10O2**	**2-Hexenoic acid**	
3.369	225.0619	C7H13O8	(2R,3R,4S,5R,6R)-2,3,4,5,6,7-Hexahydroxyheptanoic acid	

3.377	**433.1119**	**C21H21O10**	**(2S)-5-Hydroxy-2-(4-hydroxyphenyl)-4-oxo-3,4-dihydro-2H-chromen-7-yl beta-D-glucopyranoside**	Flavonoid
3.4	467.1119	C28H19O7	(6S,7S)-1,7-Bis(4-hydroxyphenyl)-6,7-dihydro-2-oxadibenzo[cd,h]azulene-4,6,8,10-tetrol	
3.406	**755.1455**	**C35H31O19**	**4-[3-(3,4-Dihydroxyphenyl)propanoyl]-3,5-dihydroxyphenyl 4,6-bis-O-(3,4,5-trihydroxybenzoyl)-beta-D-glucopyranoside**	Isoflavone
3.445	329.1138	C18H13N6O	2-Amino-7-phenyl-3-[(E)-phenyldiazenyl]pyrazolo[1,5-a]pyrimidin-5(4H)-one	
3.495	431.1335	C23H19N4O5	(2S,4R,9a'S)-1'-Hydroxy-2',2'-dimethyl-4-(4-oxo-3(4H)-quinazolinyl)-1',9a'-dihydro-3H-spiro[furan-2,9'-imidazo[1,2-a]indole]-3',5(2'H,4H)-dione	
3.498	275.0512	C9H11N2O8	2,3-Diacetoxy-4-(carbamoylamino)-4-oxobutanoic acid	
3.499	295.0443	C13H11O8	Phaseolic acid	
3.527	335.0673	C18H11N2O5	Methyl 2-[(E)-(5-oxo-2-phenyl-1,3-oxazol-4(5H)-ylidene)methyl]-4H-furo[3,2-b]pyrrole-5-carboxylate	
3.543	227.078	C8H11N4O4	4-Amino-1-(2-deoxypentofuranosyl)-1,3,5-triazin-2(1H)-one	
3.553	**162.9939**	C5H5O5	**(4S)-5-Hydroxy-2,4-dioxopentanoate**	
3.553	**119.008**	C3H3O3S	**3-mercaptopyruvate**	
3.558	475.16	C25H23N4O6	1,4:3,6-Dianhydro-2-[(1,3-benzodioxol-5-ylcarbonyl)amino]-2,5-dideoxy-5-{[4-(4-methoxyphenyl)-2-pyrimidinyl]amino}-L-iditol	
3.56	229.0617	C12H9N2O3	1-Naphthyl methyl(nitroso)carbamate	
3.568	173.0386	C14H5	1,2,4,5-Tetraethynylbenzene	
3.622	**206.0325**	**C8H17NOS2**	**Dihydrolipoamide**	
3.629	441.1165	C23H21O9	1,2-Di-O-acetyl-3,5-di-O-benzoyl-L-xylofuranose	
3.639	515.188	C34H27O5	3,3',5-Tris(benzyloxy)-2-biphenylcarboxylic acid	
3.652	449.1431	C22H25O10	3-Hydroxy-5-[(2S)-7-hydroxy-3,4-dihydro-2H-chromen-2-yl]-2-methoxyphenyl beta-D-glucopyranoside	Flavonoid
3.687	439.1613	C21H27O10	Picrodendrin V	
3.705	427.1013	C22H19O9	5-(5-Acetoxy-3,7-dimethoxy-4-oxo-4H-chromen-2-yl)-2-methoxyphenyl acetate	Flavonoid
3.731	457.1501	C25H21N4O5	(12aS)-8-(1,3-Benzodioxol-5-yl)-2-[(1-methyl-1H-pyrrol-2-yl)carbonyl]-1,3,4,12a-tetrahydropyrazino[2,1-c][1,4]benzodiazepine-6,12(2H,11H)-dione	
3.811	331.0772	C24H11O2	**Dibenzo[c,pqr]tetraphene-7,14-dione**	
3.814	**685.1758**	**C33H33O16**	**6-Methoxy-2-(4-methoxyphenyl)-4-oxo-7-[(2,3,4,6-tetra-O-acetyl-beta-D-glucopyranosyl)oxy]-4H-chromen-5-yl acetate**	Flavonoid
3.816	413.0856	C21H17O9	(2R,3S,4R)-2-(3,4-Dihydroxyphenyl)-4-(2,4,6-trihydroxyphenyl)-3,5,7-chromanetriol	Flavonoid
3.823	217.0627	C12H14N2O2	N-Acetylserotonin	**Normelatonin**

3.835	**135.9966**	**C7H6NO2**	**4-aminobenzoate**	
3.836	**262.057**	**C10H8N5O4**	**1-(2,4-Dinitrophenyl)-3,5-dimethyl-1H-1,2,4-triazole**	
3.842	203.0329	C11H8O4	1,4-dihydroxy-2-naphoate	
3.846	494.0318	C19H8N7O10	(2E)-1-(2,4-Dinitrophenyl)-2-(2,4,7-trinitro-9H-fluoren-9-ylidene)hydrazine	
3.848	273.0864	C14H13N2O4	Methyl 3,5-bis(2-cyanoethoxy)benzoate	
3.856	166.9879	C8H8O4	4-Hydroxymandelate	Phenolic acid
3.872	417.1539	C22H26O8	syringaresinol	
3.881	199.0524	C8H8O6	4-Maleylacetoacetate	
3.887	361.1655	C21H21N4O2	6,7-Dimethoxy-3-methyl-1-(4-methylphenyl)-2-(1H-1,2,4-triazol-5-yl)-1,2-dihydroisoquinoline	
3.89	353.1283	C23H17N2O2	(1,3,4-Triphenyl-1H-pyrazol-5-yl)acetic acid	
3.896	377.1243	C19H21O8	Diethyl 2,2'-[(5,9-dioxo-6,7,8,9-tetrahydro-5H-benzo[7]annulene-1,4-diyl)bis(oxy)]diacetate	
3.919	501.0477	C26H13O11	5,5'-Oxybis(4-methoxy-7-oxo-7H-furo[3,2-g]chromene-6-carbaldehyde)	
3.921	381.1337	-C23H17N4O2	N-Benzoyl-N-(5-methyl-4-phenyl-1H-1,2,3-triazol-1-yl)benzamide	
3.948	239.9903	C7H2N3O7	2,4,6-Trinitrobenzaldehyde	
3.956	403.0443	C22H11O8	6,12-Dioxo-6,12-dihydrobenzo[b][1]benzofuro[2,3-f][1]benzofuran-3,9-diyl diacetate	
3.961	531.1851	C27H31O11	(1R,2R)-3-Acetoxy-1-(4-acetoxy-3,5-dimethoxyphenyl)-2-[4-(1-acetoxyethyl)phenoxy]propyl acetate	
3.965	501.1683	C28H25N2O7	(3aR,4R,5R)-2-Phenyl-5-(3,4,5-trimethoxyphenyl)-3,3a,4,5-tetrahydro-2H-[1,3]benzodioxolo[5,6-g]indazole-4-carboxylic acid	
3.966	357.0759	C22H13O5	5-Hydroxy-4-oxo-2-phenyl-4H-chromen-7-yl benzoate	
3.977	655.1642	C32H31O15	(1S,4aS,6S,7aS)-1-({6-O-[(2E)-3-(3,4-Dihydroxyphenyl)-2-propenoyl]-2-O-(4-hydroxybenzoyl)-beta-D-glucopyranosyl}oxy)-6-hydroxy-7-methylene-1,4a,5,6,7,7a-hexahydrocyclopenta[c]pyran-4-carboxylic acid	
3.988	517.1336	C25H25O12	4,8-Dihydroxy-1-naphthyl 6-O-(4-hydroxy-3,5-dimethoxybenzoyl)-beta-D-glucopyranoside	Flavoring Chemical
3.993	439.1596	C17H23N6O8	Dimethyl 6-[(dimethylamino)(4-pyridinyl)methylene]-3,9-dioxo-2,10-dioxa-4,5,7,8-tetraazaundecane-5,7-dicarboxylate	
4.009	361.1652	C21H21N4O2	6,7-Dimethoxy-3-methyl-1-(4-methylphenyl)-2-(1H-1,2,4-triazol-5-yl)-1,2-dihydroisoquinoline	Quinoline
4.022	**164.9724**	**C8H6O4**	**Phthalic acid**	
4.023	179.0093	C7H3N2O4	6-Nitro-1,3-benzoxazol-2(3H)-one	
4.025	129.0119	C5H7NO3	1-pyrroline 4-hydroxy-2-carboxylate	

4.049	465.1393	C22H25O11	[(1S,4aR,5S,7aS)-1-(beta-D-Glucopyranosyloxy)-5-hydroxy-1,4a,5,7a-tetrahydrocyclopenta[c]pyran-7-yl]methyl 4-hydroxybenzoate	
4.051	393.1177	C19H21O9	1,2,3-Tri-O-acetyl-4,6-O-benzylidene-beta-D-mannopyranose	
4.061	215.0463	C11H7N2O3	5-Benzylidene-2,4,6(1H,3H,5H)-pyrimidinetrione	
4.062	187.0538	C9H16O4	Azelaic acid	
4.109	524.1427	C23H26NO13	7-Methoxy-3-oxo-3,4-dihydro-2H-1,4-benzoxazin-2-yl 2,3,4,6-tetra-O-acetyl-beta-D-glucopyranoside	
4.11	591.1658	C35H27O9	(5S,6R,7S,8S)-4-Oxaspiro[2.5]octane-5,6,7,8-tetrayl tetrabenzoate	
4.121	243.0758	C13H11N2O3	5-Allyl-5-phenyl-2,4,6(1H,3H,5H)-pyrimidinetrione	
4.135	559.1446	C27H27O13	(1S,3R,4R,5R)-3-{[(2E)-3-(3,4-Dihydroxyphenyl)-2-propenoyl]oxy}-1,5-dihydroxy-4-{[(2E)-3-(4-hydroxy-3,5-dimethoxyphenyl)-2-propenoyl]oxy}cyclohexanecarboxylic acid	
4.138	309.0736	C13H13N2O7	5-[(3S,4S,5S)-5-Carboxy-4-(carboxymethyl)-3-pyrrolidinyl]-6-oxo-1,6-dihydro-2-pyridinecarboxylic acid	
4.143	317.1492	C17H21N2O4	1-(4-Methoxyphenoxy)-3,3-dimethyl-2-butanyl 1H-imidazole-1-carboxylate	
4.155	685.2336	C45H29N6O2	{2-[(E)-(8-{[2-(2-Aminobenzoyl)phenyl]amino}-6H-quinolino[2,3,4-kl]acridin-6-ylidene)amino]phenyl}(2-aminophenyl)methanone	
4.173	**171.059**	**C7H7O5**	**3-Dehydroshikimate**	
4.211	150.9947	C8H7O3	2-Hydroxyphenylacetate	
4.242	552.2131	C32H30N3O6	(1S)-1,4-Anhydro-2,3,5-tri-O-benzyl-1-[5-cyano-4-(methoxycarbonyl)-1H-pyrazol-3-yl]-D-ribitol	
4.301	373.1273	C20H21O7	4-[(2,4-Dimethoxy-3,6-dimethylbenzoyl)oxy]-2-hydroxy-3,6-dimethylbenzoic acid	
4.342	326.1021	C18H16NO5	Methyl (2E)-4-[(4-methoxyfuro[2,3-b]quinolin-7-yl)oxy]-2-methyl-2-butenoate	
4.427	305.0172	C11H5N4O7	1-(2,4,6-Trinitrophenyl)-4(1H)-pyridinone	
4.439	655.2008	C28H35N2O16	2,5-Pyrazinediylbis(1R,2S,3R)-3,4-diacetoxy-1,1,2-butanetriyl tetraacetate	
4.454	136.9792	C7H5O3	Hydroxybenzoate	
4.47	585.2156	C22H37N2O16	beta-D-Mannopyranosyl-(1->4)-2-acetamido-2-deoxy-beta-D-glucopyranosyl-(1->4)-2-acetamido-2-deoxy-beta-D-glucopyranose	
4.501	**193.9967**	**C9H9NO4**	**alpha-Hydroxybenzoylglycine**	
4.518	225.1005	C11H17NO4	(4-aminophenyl)-1,2,3,4-tetrahydroxypentane	
4.518	**329.103**	**C18H17O6**	**1-(5,8-Dihydroxy-1,4-dioxo-1,4-dihydro-2-naphthalenyl)-4-methyl-3-penten-1-yl acetate**	
4.532	343.1544	C21H19N4O	(2E)-3-[4-(Dimethylamino)phenyl]-2-(6-ethyl-4-oxo-1,4-dihydro-2-quinazolinyl)acrylonitrile	
4.539	231.1087	C8H15N4O4	2,2'-(1,4-Piperazinediyl)bis(N-hydroxyacetamide)	

4.6	239.0762	C8H15O8	(5S)-5-[(1R,2R)-1,2,3-Trihydroxypropyl]-alpha-D-xylulofuranose	
4.6	327.0502	C17H11O7	5-Methoxy-3,4,7a,10a-tetrahydro-1H,12H-furo[3',2':4,5]furo[2,3-h]pyrano[3,4-c]chromene-1,12-dione	
4.601	305.0459	C18H9O5	3,5-Dihydroxy-1H,3H-isochromeno[6,5,4-mna]xanthen-1-one	
4.627	193.002		3-Dehydrogulonate	
4.691	201.0672	C11H9N2O2	5-Methyl-1,3,4,5-tetrahydropyrrolo[4,3,2-de]quinoline-7,8-dione	
4.699	809.2272	C40H41O18	3,6-Bis-O-[(2E)-3-(4-hydroxyphenyl)-2-propenoyl]-beta-D-fructofuranosyl 6-O-[(2E)-3-(4-hydroxy-3-methoxyphenyl)-2-propenoyl]-alpha-D-glucopyranoside	Flavonoid
4.699	373.1276	C20H21O7	4-[(2,4-Dimethoxy-3,6-dimethylbenzoyl)oxy]-2-hydroxy-3,6-dimethylbenzoic acid	
4.743	421.15	C21H25O9	(1S,3S)-7,10-Dihydroxy-1,3-dimethyl-3,4-dihydro-1H-benzo[g]isochromen-9-yl beta-D-glucopyranoside	Flavonoid
4.792	**201.0679**	**C10H18O4**	**Sebacic acid**	
4.816	751.2253	C38H39O16	2-[(2R)-7-Oxo-2,3-dihydro-7H-furo[3,2-g]chromen-2-yl]-2-propanyl 6-O-[(2E)-3-(4-acetoxy-3-methoxyphenyl)-2-propenoyl]-2,3,4-tri-O-acetyl-beta-D-glucopyranoside	Flavonoid
4.818	371.1148	C20H19O7	1,3,6,8-Tetrahydroxy-2-(1-hydroxyhexyl)-9,10-anthraquinone	
4.895	343.1602	C14H23N4O6	N,N'-[(3,6-Dioxo-2,5-piperazinediyl)di-3,1-propanediyl]bis(N-hydroxyacetamide)	
4.897	265.0244	C14H5N2O4	Benzo[lmn][3,8]phenanthroline-1,3,6,8(2H,7H)-tetrone	
4.91	685.1741	C33H33O16	6-Methoxy-2-(4-methoxyphenyl)-4-oxo-7-[(2,3,4,6-tetra-O-acetyl-beta-D-glucopyranosyl)oxy]-4H-chromen-5-yl acetate	
4.919	255.0169	C14H8O5	1,2,4-Trihydroxy-9,10-anthraquinone	Quinone
4.919	493.169	C24H29O11	2-Hydroxy-3-methoxy-6-[(Z)-2-(3,4,5-trimethoxyphenyl)vinyl]phenyl beta-D-glucopyranoside	
4.932	469.1704	C22H29O11	(4R)-6-(Benzyloxy)-4-({[(2R,3R)-1,4-diethoxy-3-hydroxy-1,4-dioxo-2-butanyl]oxy}carbonyl)-4-hydroxyhexanoic acid	
4.966	595.2211	C38H31N2O5	N~2~-[(9H-Fluoren-9-ylmethoxy)carbonyl]-N-trityl-L-asparagine	
4.982	675.7427	C46H92O2	Hexatetracontanoic acid	
5.027	565.1885	C22H33N2O15	beta-D-Glucopyranosyl-(1->4)-beta-D-glucopyranosyl-(1->5)-2'-deoxy-3,4-dihydrothymidine	
5.06	645.7292	C46H93	Hexatetracontane	

5.102	643.7296	C46H92	Tetracontylcyclohexane	
5.104	415.1369	C18H19N6O6	9-(2-Deoxy-beta-D-erythro-pentofuranosyl)-6-[2-(4-nitrophenyl)ethoxy]-9H-purin-2-amine	
5.121	433.1577	C33H21O	(1E,4E)-1,5-Di(9-anthryl)-1,4-pentadien-3-one	
5.13	660.2367	C38H34N3O8	1-{3-(Acetoxymethyl)-2,3-dideoxy-5-O-[(4-methoxyphenyl)(diphenyl)methyl]-beta-D-threo-pentofuranosyl}-4-(2-nitrophenyl)-2(1H)-pyrimidinone	
5.193	645.2487	C28H41N2O15	Methyl 2-acetamido-3,6-di-O-acetyl-2,4-dideoxy-4-{[(1R,2S,3S,4S,5R)-2,3,4-triacetoxy-5-(acetoxymethyl)cyclohexyl]amino}-beta-D-glucopyranoside	
5.194	602.2109	C26H36NO15	Methyl (1R,2S,3S,5S)-2-acetamido-5-acetoxy-3-[(2,3,4,6-tetra-O-acetyl-beta-D-glucopyranosyl)oxy]cyclohexanecarboxylate	
5.332	649.2296	C35H37O12	1,2,3,4-Tetra-O-acetyl-6-O-[bis(4-methoxyphenyl)(phenyl)methyl]-D-mannopyranose	
5.353	357.2033	C18H25N6O2	(3R,6R)-N-(2-Furylmethyl)-6-({4-[(methylamino)methyl]-1H-1,2,3-triazol-1-yl}methyl)quinuclidine-3-carboxamide	
5.368	681.2415	C32H41O16	4-{(1S,3aR,4S,6aR)-4-[4-(beta-D-Glucopyranosyloxy)-3-methoxyphenyl]tetrahydro-1H,3H-furo[3,4-c]furan-1-yl}-2-methoxyphenyl beta-D-glucopyranoside	
5.386	627.2557	C45H31N4	21-Methyl-5,10,15,20-tetraphenylporphyrin	
5.395	152.9741	C7H6O4	2,3-Dihydroxybenzoic acid	
5.497	327.1603	C21H19N4	[2-(4-Benzyl-1-piperazinyl)benzylidene]malononitrile	
5.5	719.2481	C37H39N2O13	(2R,3R,3aS,5R,6R,7S,7aS)-3-Acetoxy-3a-(benzyloxy)-7a-[(benzyloxy)methyl]-2-(5,6-dihydroxy-5-methyl-2,4-dioxotetrahydro-1(2H)-pyrimidinyl)-5,6-dihydroxyoctahydro-1-benzofuran-7-yl benzoate	
5.573	242.1274	C9H13N3O5	Cytidine	
5.614	651.3477	C34H47N6O7	Methyl {(2S,6R)-4-[(2S)-1-amino-1-oxo-3-phenyl-2-propanyl]-6-[(1S)-5-{[(2-methylphenyl)carbamoyl]amino}-1-({[(2-methyl-2-propanyl)oxy]carbonyl}amino)pentyl]-3-oxo-2-piperazinyl}acetate	
5.616	379.1148	C17H19N2O8	Diethyl 4-hydroxy-1-(4-nitrobenzoyl)-2,2-pyrrolidinedicarboxylate	
5.764	811.3865	C51H55O9	(2S)-6-({(2S)-5,7-Dihydroxy-2-(4-hydroxyphenyl)-8-[(2R)-2-isopropenyl-5-methyl-5-hexen-1-yl]-4-oxo-3,4-dihydro-2H-chromen-6-yl}methyl)-5,7-dihydroxy-8-[(2R)-2-isopropenyl-5-methyl-5-hexen-1-yl]-2-phen	Flavonoid
5.782	**183.0901**	**C9H12O4**	**3-methoxy-4-hydroxyphenylglycol**	
5.784	397.16	C18H25N2O8	3-(4-Ethoxy-4-oxobutyl)-2',3'-O-isopropylideneuridine	

5.803	229.0975	C13H13N2O2	Methyl 1-(1-phenylethyl)-1H-imidazole-5-carboxylate	
5.821	169.0392		3,4-Dihydroxyphenylglycol	
5.91	213.1022	C13H13N2O	9-Amino-1,2,3,4-tetrahydro-1-acridinol	
6.124	**243.1119**	**C10H15N2O3S**	**Biotin**	**Vitamin**
6.17	255.0763	C14H11N2O3	N-(4-Nitrophenyl)-2-phenylacetamide	
6.216	243.1125	C14H15N2O2	3,3'-Dimethoxy-4,4'-biphenyldiamine	
6.218	528.1953	C20H34NO15	2-Acetamido-2-deoxy-alpha-D-mannopyranosyl-(1->4)-alpha-D-glucopyranosyl-(1->2)-6-deoxy-alpha-L-mannopyranose	
6.26	453.2206	C17H33N4O10	(2R,3S,4R,5R,6R)-5-amino-2-(aminomethyl)-6-[(1R,2R,4R,6S)-4,6-diamino-2-[(2S,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl]oxy-3-hydroxy-cyclohexoxy]tetrahydropyran-3,4-diol	
6.286	301.1524	C12H21N4O5	Alanylalanylalanylalanine	
6.294	503.278	C27H39N2O7	Ethyl (2-{[(2R)-2-(2,5-dimethoxyphenyl)-2-{[(1R,2R,4S,6R,7R)-1,10,10-trimethyl-3-oxatricyclo[5.2.1.0~2,6~]dec-4-yl]oxy}ethyl]amino}-2-oxoethyl)carbamate	
6.397	513.1898	C31H29O7	Bis(3',4',5'-trimethoxy-4-biphenylyl)methanone	
6.417	528.2009	C31H30NO7	Ethyl 2-{[1-formyl-6,7-dimethoxy-2-(4-methylbenzoyl)-1,2-dihydro-1-isoquinolinyl]methyl}-6-methoxybenzoate	
6.458	313.1509	C13H21N4O5	[(2S,6S)-6-(Azidomethyl)-4-{2-[(2-methyl-2-propanyl)oxy]-2-oxoethyl}-2-morpholinyl]acetic acid	
6.491	227.1168	C14H12O3	Resveratrol	
6.512	275.1298	C17H15N4	2,2'-Methylenebis(1-methyl-1H-benzimidazole)	
6.703	826.4947	C44H68N5O10	4-[(1-Hydroxy-2-oxo-3-azepanyl)amino]-4-oxo-2-butanyl N~6~-hydroxy-N~6~-[(2E)-2-octadecenoyl]-N~2~-{[(2E)-2-(6-oxo-2,4-cyclohexadien-1-ylidene)-1,3-oxazolidin-4-yl]carbonyl}lysinate	
6.727	325.1448	C20H21O4	5,5'-Diallyl-3,3'-dimethoxy-2,2'-biphenyldiol	
6.753	257.1292	C15H17N2O2	Propyl 1-(1-phenylethyl)-1H-imidazole-5-carboxylate	
6.864	595.2304	C31H35N2O10	[(1R,2R,3S,4S,5R,6S)-2,4-Diacetoxy-3-(benzoyloxy)-5-(cyclohexylamino)-6-nitrocyclohexyl]methyl benzoate	
6.873	241.1329	C15H17N2O	1-(4-Amino-2-methyl-5-phenyl-1H-pyrrol-3-yl)-2-methyl-1-propanone	
6.89	909.4521	C46H69O18	(2R,3R,4aR,4bS,7R,8aR)-3-Hydroxy-4a-methyl-7-(2-methyl-3-furyl)-8-oxo-1,2,3,4,4a,4b,5,6,7,8,8a,9-dodecahydro-2-phenanthrenyl beta-D-glucopyranosyl-(1->4)-2,6-dideoxy-3-O-methyl-alpha-L-ribo-hexopyrano	

6.931	311.1652	C20H23O3	(17beta)-3-Oxoestra-4,9,11-trien-17-yl acetate	
6.999	483.21	C20H35O13	beta-D-Glucopyranosyl-(1->4)-6-deoxy-3-O-methyl-beta-D-allopyranosyl-(1->4)-2,6-dideoxy-3-O-methyl-beta-D-arabino-hexopyranose	
7.037	779.418	C48H59O9	(3beta)-27-{[(2E)-3-(3,4-Dihydroxyphenyl)-2-propenoyl]oxy}-3-{[(2E)-3-(4-hydroxyphenyl)-2-propenoyl]oxy}olean-12-en-28-oic acid	
7.074	597.2529	C29H41O13	(3E)-4-{(1S,4S,6R)-2,2,6-Trimethyl-4-[(2,3,4,6-tetra-O-acetyl-beta-D-glucopyranosyl)oxy]-7-oxabicyclo[4.1.0]hept-1-yl}-3-buten-2-yl acetate	
7.111	593.206	C25H37O16	Benzyl beta-D-galactopyranosyl-(1->2)-[beta-D-galactopyranosyl-(1->3)]-beta-D-galactopyranoside	
7.164	879.3709	C54H55O11	2-O-Benzoyl-4-O-benzyl-6-deoxy-3-O-(2,3,4,6-tetra-O-benzyl-alpha-D-glucopyranosyl)-alpha-L-mannopyranose	
7.171	339.1595	C22H19N4	(Z)-2-(5,6-Dimethyl-1H-benzotriazol-1-yl)-1,2-diphenylethenamine	
7.172	194.033	C7H9N5O2	2-amino-4-hydroxy-6-hydroxy-methyl-dihydropteridine	
7.218	633.3375	C34H45N6O6	1,1'-(1,4,10,13-Tetraoxa-7,16-diazacyclooctadecane-7,16-diyl)bis[2-amino-3-(1H-indol-3-yl)-1-propanone]	
7.228	595.2351	C36H35O8	(3'R)-2,3'-Bis(2,4-dimethoxyphenyl)-7,7'-dimethoxy-3',4'-dihydro-2'H,4H-3,6'-bichromene	
7.24	313.1817	C20H25O3	7-Oxoabieta-8,11,13-trien-18-oic acid	
7.248	597.2478	C37H33N4O4	[(4S,5S)-2,2-Dimethyl-4,5-diphenyl-1,3-imidazolidinediyl]bis{[(5S)-3-phenyl-4,5-dihydro-1,2-oxazol-5-yl]methanone}	
7.27	285.1572	C13H17N8	4-(1H-Imidazol-1-ylmethyl)-6-(3,5,5-trimethyl-4,5-dihydro-1H-pyrazol-1-yl)-1,3,5-triazin-2-amine	
7.281	822.4637	C43H68NO14	(3beta,16alpha,21beta)-3-{[2-Acetamido-6-O-(alpha-L-arabinopyranosyl)-2-deoxy-beta-D-glucopyranosyl]oxy}-16,21-dihydroxyolean-12-en-28-oic acid	
7.283	201.0627	C6H9N4O4	1-Methoxy-3-(3-nitro-1H-1,2,4-triazol-1-yl)-2-propanol	
7.302	313.1812	C20H25O3	7-Oxoabieta-8,11,13-trien-18-oic acid	
7.372	399.2229	C18H31N4O6	N-Acetylisoleucylleucyl-alpha-asparagine	
7.391	355.1973	C16H27N4O5	N~2~-{{(3R,4R,5R)-4,5-Dihydroxy-3-[(N-methylglycyl)amino]-1-cyclohexen-1-yl}carbonyl)-D-leucinamide	
7.442	313.1866	C14H25N4O4	N'~1~,N'~4~-Bis(2,2-dimethylpropanoyl)succinohydrazide	
7.474	507.202	C31H23N8	2,2'-(2-Methyl-1,4-phenylene)bis{[4-(phenyldiazenyl)phenyl]diazene}	
7.5	299.1725	C13H23N4O4	N'~1~,N'~5~-Diisobutyrylpentanedihydrazide	

7.549	361.1397	C18H21N2O6	(17beta)-2,4-Dinitroestra-1(10),2,4-triene-3,17-diol	
7.559	357.2131	C16H29N4O5	L-Alanyl-N~1~-(6-acetoxyhexyl)-D-glutamamide	
7.58	483.2046	C28H27N4O4	2,5-Bis[(1R,9S)-6-oxo-7,11-diazatricyclo[7.3.1.0~2,7~]trideca-2,4-dien-11-yl]-1,4-benzoquinone	
7.586	369.2158	C18H25N8O	1-{1-[4,6-Di(1-piperidinyl)-1,3,5-triazin-2-yl]-5-methyl-1H-1,2,3-triazol-4-yl}ethanone	
7.591	285.1578	C13H17N8	4-(1H-Imidazol-1-ylmethyl)-6-(3,5,5-trimethyl-4,5-dihydro-1H-pyrazol-1-yl)-1,3,5-triazin-2-amine	
7.592	339.1572	C17H19N6O2	6-{[(2,5-Dimethoxyphenyl)(methyl)amino]methyl}pyrido[3,2-d]pyrimidine-2,4-diamine	
7.602	297.1934	C14H25N4O3	L-Prolyl-N-methyl-D-leucylglycinamide	
7.635	826.4975	C44H68N5O10	4-[(1-Hydroxy-2-oxo-3-azepanyl)amino]-4-oxo-2-butanyl N~6~-hydroxy-N~6~-[(2E)-2-octadecenoyl]-N~2~-{[(2E)-2-(6-oxo-2,4-cyclohexadien-1-ylidene)-1,3-oxazolidin-4-yl]carbonyl}lysinate	
7.635	325.1857	-C14H29O8	(5S,6R,7S,8S,9S,10S,11S)-1,5,6,7,8,9,10,11-Tetradecaneoctol	
7.667	981.4662	C49H73O20	(3beta,22beta)-24-Hydroxy-29-oxo-22,29-epoxyolean-12-en-3-yl methyl 6-deoxy-alpha-L-mannopyranosyl-(1->2)-beta-D-glucopyranuronosyl-(1->2)-beta-D-glucopyranosiduronate	
7.676	369.2146	C17H29N4O5	Valylprolylglycylvaline	
7.7	378.2427	C25H32NO2	2,2-Dimethyl-7-(3-methyl-2-octanyl)-4-(4-pyridinyl)-2H-chromen-5-ol	
7.718	492.2091	C34H26N3O	Phenyl(1,2,4,5-tetraphenyl-1,2,3,4-tetrahydro-1,2,4-triazin-3-yl)methanone	
7.736	452.2182	C25H30N3O5	(2E,4E)-N-[(1S)-3-({(2S)-3-Methyl-1-[(3R,4S)-4-methyl-2,5-dioxo-3-pyrrolidinyl]-1-oxo-2-butanyl}amino)-3-oxo-1-phenylpropyl]-2,4-hexadienamide	
7.756	769.4392	C40H65O14	(1beta,3alpha,5alpha,6beta,22R,25S)-26-(beta-D-Glucopyranosyloxy)-22-hydroxy-6-methoxy-3,5-cyclofurostan-1-yl 6-deoxy-beta-D-galactopyranoside	
7.764	311.1181	C21H15N2O	4-(4-Methoxyphenyl)-2-phenylquinazoline	
7.772	267.1455	C12H19N4O3	Leucylhistidine	
7.782	221.104	C10H13N4O2	3-Isobutyl-1-methyl-3,7-dihydro-1H-purine-2,6-dione	
7.801	283.1773	C13H23N4O3	L-Prolyl-L-leucylglycinamide	
7.808	363.1554	C19H19N6O2	3,7-Dimethyl-1-[4-(1,8-naphthyridin-2-yl)butyl]-3,7-dihydro-1H-purine-2,6-dione	
7.819	323.1734	C15H23N4O4	2-Methyl-2-propanyl (2R,4R)-2-(aminomethyl)-4-(5-methyl-2,4-dioxo-3,4-dihydro-1(2H)-pyrimidinyl)-1-pyrrolidinecarboxylate	

7.843	452.2178	C25H30N3O5	(2E,4E)-N-[(1S)-3-({(2S)-3-Methyl-1-[(3R,4S)-4-methyl-2,5-dioxo-3-pyrrolidinyl]-1-oxo-2-butanyl}amino)-3-oxo-1-phenylpropyl]-2,4-hexadienamide	
7.853	397.2249	C22H29N4O3	2-[4-(Diethylcarbamoyl)-6-methyl-2-pyridinyl]-N,N-diethyl-6-methylisonicotinamide 1-oxide	
7.892	507.2041	C31H23N8	2,2'-(2-Methyl-1,4-phenylene)bis{[4-(phenyldiazenyl)phenyl]diazene}	
7.895	483.2082	C20H35O13	beta-D-Glucopyranosyl-(1->4)-6-deoxy-3-O-methyl-beta-D-allopyranosyl-(1->4)-2,6-dideoxy-3-O-methyl-beta-D-arabino-hexopyranose	
7.919	857.4509	C43H69O17	(3beta,16beta)-3,17-Dihydroxy-22-oxocholest-5-en-16-yl beta-D-glucopyranosyl-(1->4)-beta-D-xylopyranosyl-(1->3)-alpha-L-arabinopyranoside	
7.919	295.1702	C20H23O2	ethynylestradiol	
7.933	383.2591	C25H35O3	(17beta)-3-Hydroxyestra-1,3,5(10)-trien-17-yl heptanoate	
7.936	714.4459	C37H64NO12	(2R,3R,6R,7S,8S,9R,10R)-3-[(2R,3R)-2,3-Dihydroxy-2-pentanyl]-9-{[(2S,3R,4S,6R)-4-(dimethylamino)-3-hydroxy-6-methyltetrahydro-2H-pyran-2-yl]oxy}-7-{[(2R,4R,5S,6S)-5-hydroxy-4-methoxy-4,6-dimethyltetra	
7.948	325.1337	C22H17N2O	(3,5-Diphenyl-4,5-dihydro-1H-pyrazol-1-yl)(phenyl)methanone	
7.954	738.4426	C39H64NO12	(2R,3R,4S,5R,8R,9S,10S,11R,12R)-5-Ethyl-3,4-dihydroxy-9-{[(2R,4R,5S,6S)-5-hydroxy-4-methoxy-4,6-dimethyltetrahydro-2H-pyran-2-yl]oxy}-11-({(2S,3R,4S,6R)-3-hydroxy-6-methyl-4-[methyl(2-propyn-1-yl)amin	
7.956	625.2829	C31H45O13	(1S,2R,3R,4S,5S,6S)-2,3,5-Trihydroxy-4,6-dimethoxycyclohexyl 4-(beta-D-glucopyranosyloxy)-3,5-bis(3-methyl-2-buten-1-yl)benzoate	
7.959	327.2022	C15H27N4O4	Diethyl 4,4'-methylenedi(1-piperazinecarboxylate)	
7.962	311.1754	C19H23N2O2	2-(Cyclohexylcarbonyl)-1,2,3,6,7,11b-hexahydro-4H-pyrazino[2,1-a]isoquinolin-4-one	
7.973	393.2174	C24H29N2O3	Carfentanyl	
7.978	566.284	C31H40N3O7	(3beta,16beta,20E)-17-Hydroxy-20-[(4-nitrophenyl)hydrazono]pregn-5-ene-3,16-diyl diacetate	Terpenoid

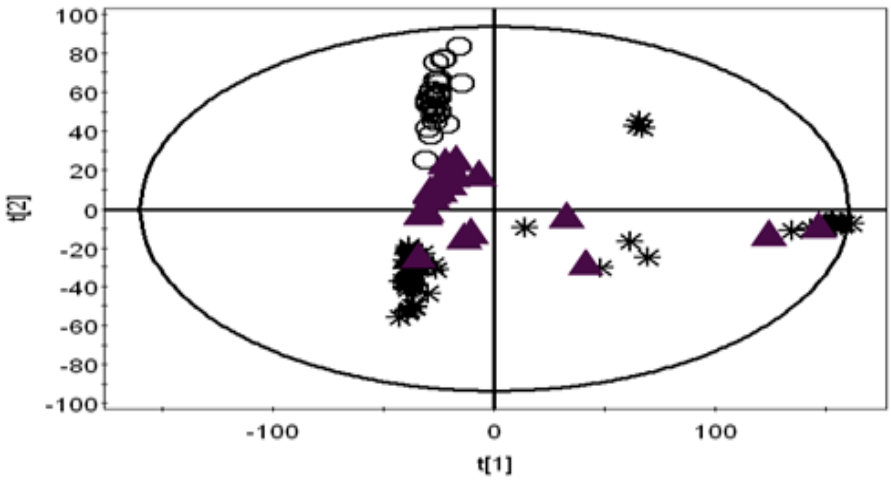
Results and Discussion - Objective 2.1.: Compound / constituent identification in finished product

Statistical analysis methodology was developed and finalized to perform non-targeted, compositional testing and validated using 26 milk, four miso, 10 natto, two speed and eight tofu samples. Training was provided to Mark MacDuff of Sevita International to perform statistical analysis using statistical methodology on MarkerLynx.

PCA and discriminate analysis of soy seed and soymilk samples (Lot 2):

Raw seeds of nine soybean varieties were compared with 12 soymilk samples (each injected in triplicate). Scores plots of principal component analysis of the full metabolome indicate that raw soybean seed and soymilk samples share considerable similarity in chemical composition (Figure 4) due to the close proximity of points.

Figure 4: PCA analysis of non-soymilk check seeds (0), soymilk (*) and soymilk seed (▲)



In order to find unique biomarkers that differentiate the raw seed and finished soymilk products, discriminant analyses were performed using an S plot (Figure 5) which compares all soymilk analysis to the seeds used in the soymilk production (soymilk seeds).

Figure 6 shows the distribution of the soymilk and soybean seed biomarkers plotted across all samples. The results in Table 2 show the identification of unique biomarkers in seeds and soymilk from the discriminant analysis identified by retention time (rt) and accurate mass. Marker # 5 (Rt 4.67 minutes) is enhanced 9.7 times in milk and is a reliable marker for soymilk quality. It is an unknown compound and needs to be isolated for further characterization and identification.
The plot of occurrence of these markers in the seed and soymilk samples shows that they are reliable biomarkers which appear in either seed or soymilk samples, but not both.

Figure 5: Plot of soybean seed versus soymilk

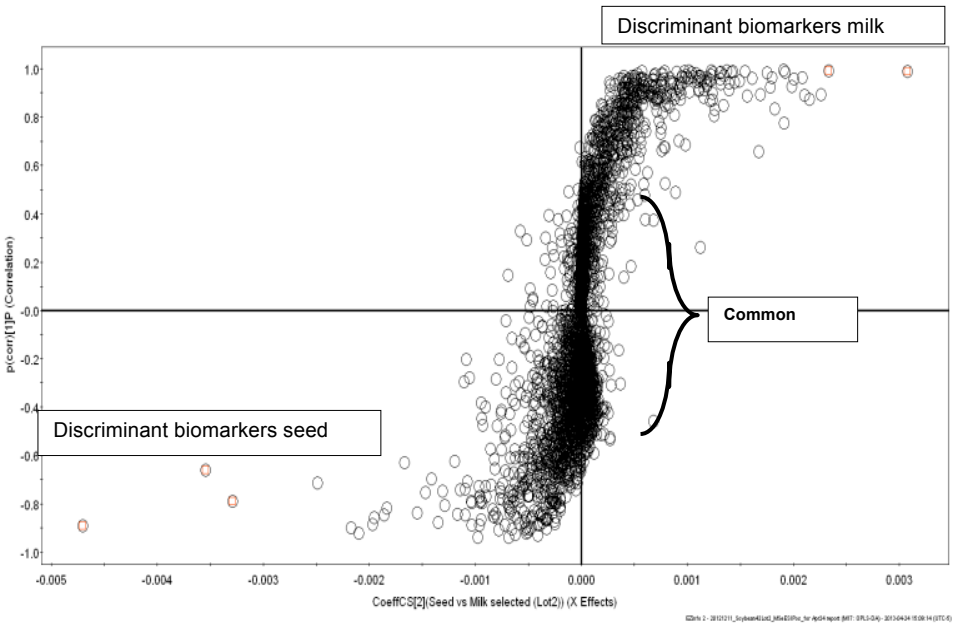
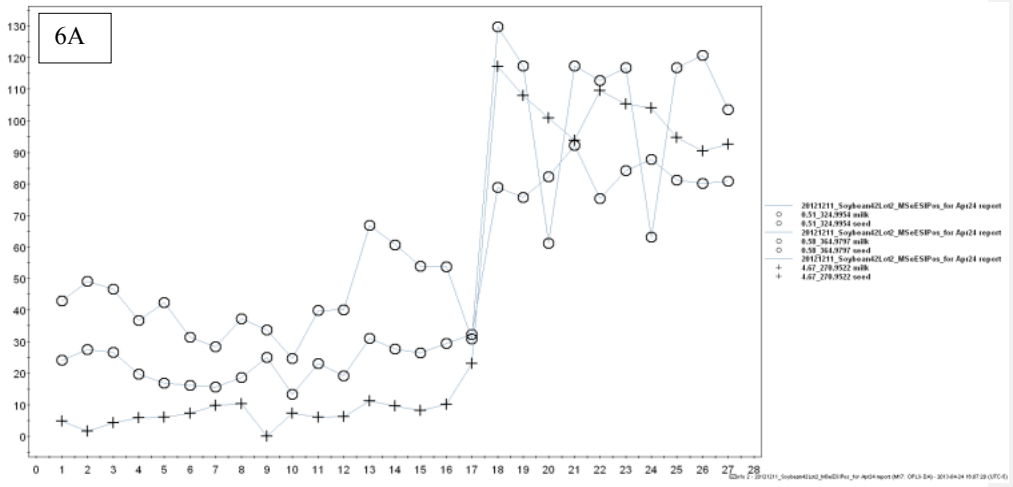


Figure 6: Distribution of soymilk (A) and soybean seed (B) biomarkers plotted across all samples



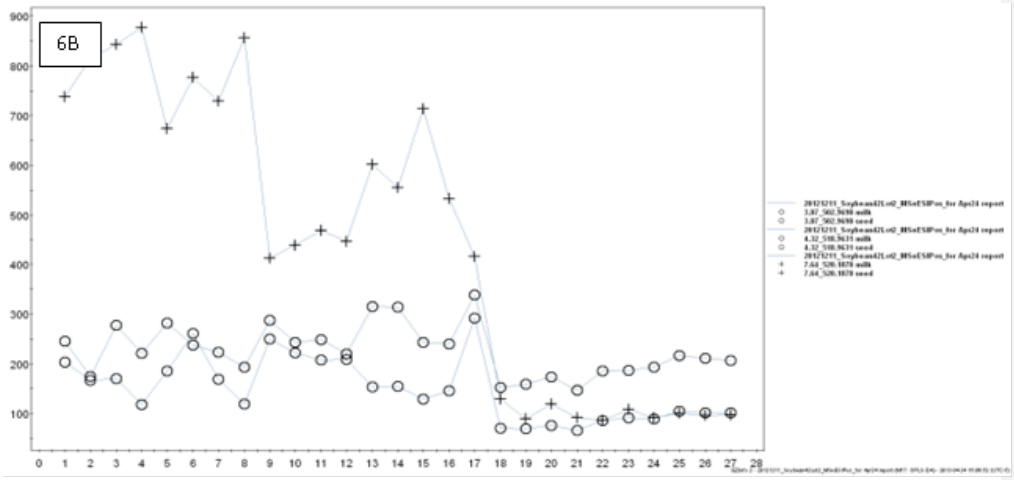


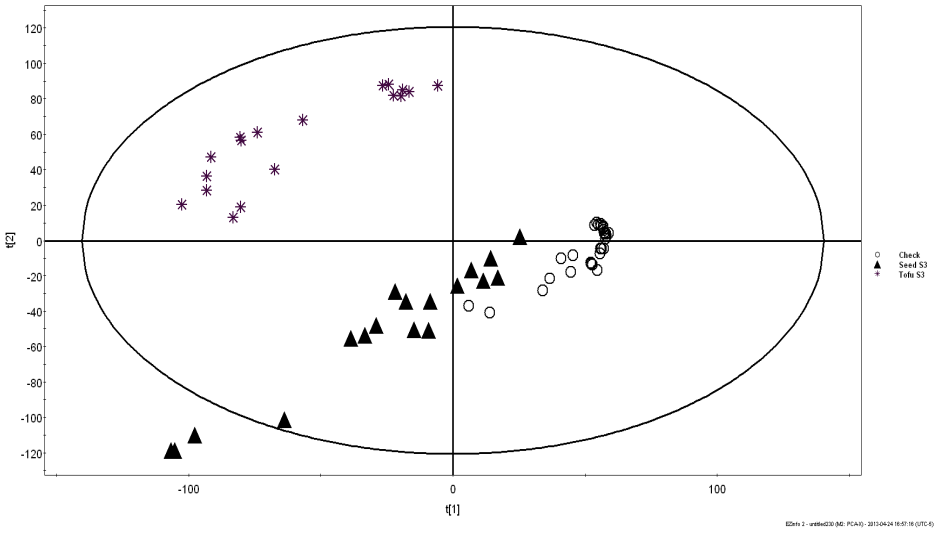
Table 2: Discriminant biomarkers of seed and soymilk

Primary ID	Retention time (min.)	Mass (atomic mass unit)	p[1]P	p(corr)[1]P	Seed	Milk	Factor of Change	Uncertainty
0.58_364.9797	0.58	364.9797	0.116863	0.810393	42.0581	113.123	2.7	0.025
0.47_705.0138	0.47	705.0138	0.1245	0.703794	58.4342	155.758	2.7	0.039
3.30_254.9593	3.3	254.9593	0.124824	0.649297	180.02	281.064	1.6	0.048
3.79_270.9512	3.79	270.9512	0.174032	0.828167	235.783	390.592	1.7	0.041
4.67_270.9522	4.67	270.9522	0.11888	0.862411	8.07613	78.1927	9.7	0.015
7.64_520.1878	7.64	520.1878	-0.30029	-0.834238	629.788	161.705	3.9	0.271

PCA analysis and discriminant analysis of raw soybean seed and tofu samples

Samples extracted from six soybean varieties' seeds were compared with samples extracted from six tofu types (each injected in triplicate). Scores plots of principal component analysis of the metabolome data indicate that seed and tofu samples are chemically different and plot in distinct groups. This is a different result as compared to soybean seed and soymilk comparison, where no separation was observed. Figure 7 shows clear clustering of seed and tofu while check samples are cluster close to the raw seed.

Figure 7: PCA analysis of non-tofu check seeds (0), tofu (*) and tofu seed (▲)



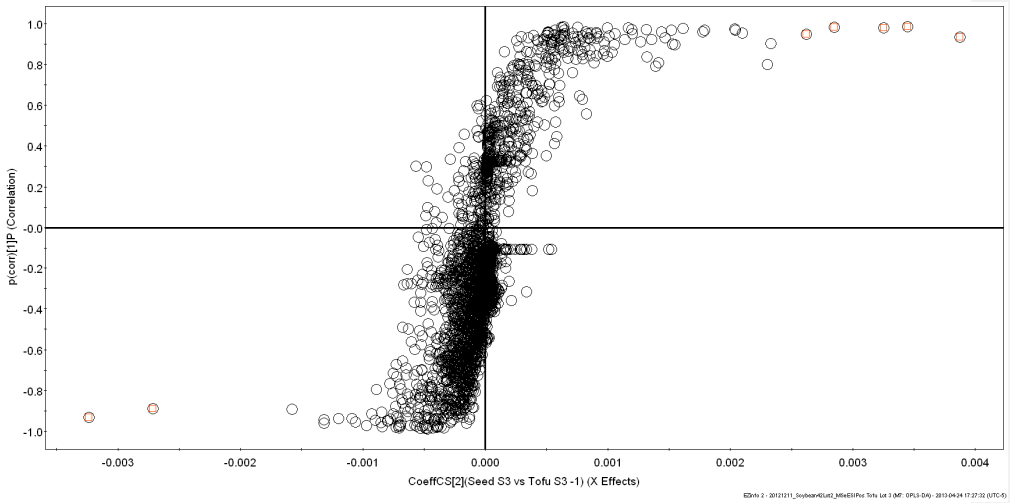
Tofu samples showed two clusters indicating two categories of tofu products (Figure 8). Check samples were excluded to perform discriminant analysis. Biomarkers were found that clearly are present in either seed or tofu and are reliable indicators across all samples (Figure 8) and their distribution is presented in Figure 9.

Table 3 shows the identification of a biomarker at Rt 6.61 min and m/z 943.3197 with 7.9 times higher occurrence in tofu than seed which is a reliable tofu biomarker.

Table 3. Discriminant biomarkers of tofu and seed

Primary ID	Retention time (min.)	Mass (atomic mass unit)	Seed	Tofu	Factor of change	Uncertainty
6.61_943.3197	6.61	943.3197	107.332	845.394	7.9	0.048
2.53_295.0109	2.53	295.0109	264.945	56.6018	4.7	0.256
1.35_311.0026	1.35	311.0026	159.025	24.3278	6.5	0.709

Figure 8. Discriminant analysis of tofu seed (left) versus tofu (right)



Conclusions and Next Steps

Six biomarkers soymilk and three biomarkers for tofu have been discovered.

Compounds found in both the raw soybean seed and the end products can now be used to screen Sevita International's germplasm to determine if there are any particular varieties that can be identified for end user testing and future sensory evaluation studies.

References

Berhow MA and Cantrell CL (2002) Analysis and Quantitative Determination of Group B Saponins in Processed Soybean Products Phytochem. Anal. 13, 343–348.

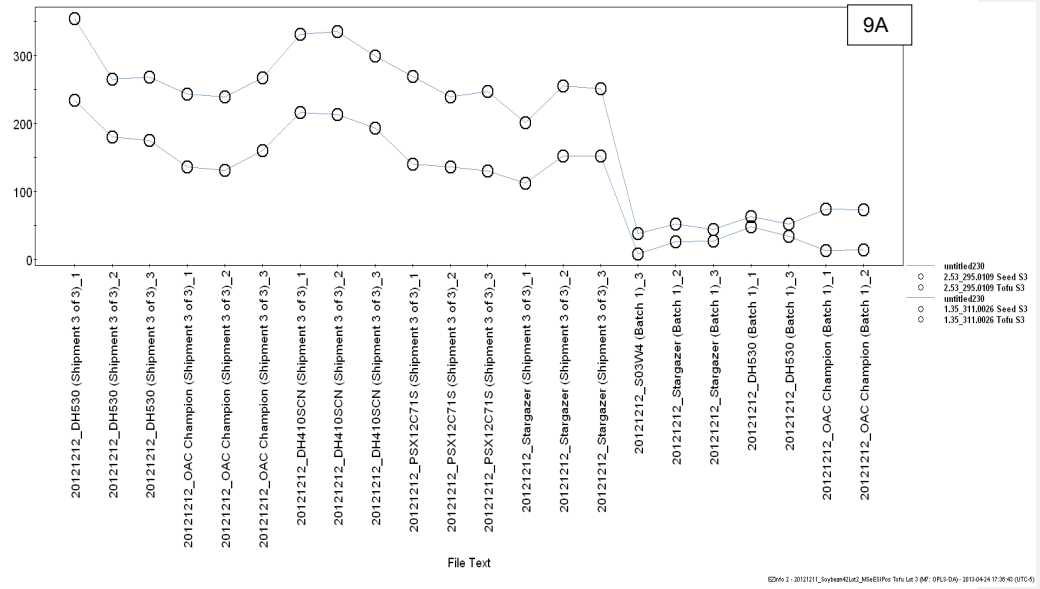
Kaneko S, Kumazawa K and Nishimura O (2011). Studies on the Key Aroma Compounds in Soy Milk Made from Three Different Soybean Cultivars. J. Agric. Food Chem. 2011, 59, 12204–12209.

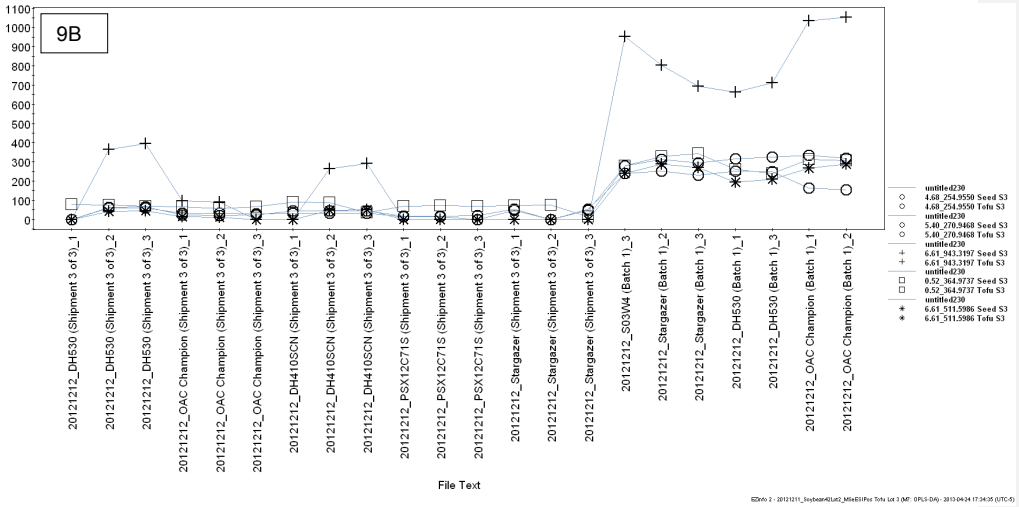
Kovinich N, Arnason JT, DeLuca V and Miki B (2011) Coloring Soybeans with Anthocyanins? In DR Gang (ed.), Recent Advances in Phytochemistry, Vol 41. Springer, pp 47-57.

Kovinich N, Saleem A, Arnason JT and Miki B (2010) Functional characterization of a UDP-glucose:flavonoid 3-O-glucosyltransferase from the seed coat of black soybean (*Glycine max* (L.) Merr.). Phytochemistry 71: 1253-1263.

Lee EJ, Shaykhutdinov R, Weljie AM, Vogel HJ, Facchini PJ, Park SU, Kim YK and Yang TJ (2009) Quality Assessment of Ginseng by 1H NMR Metabolite Fingerprinting and Profiling Analysis, J. Agric. Food Chem. 2009, 57, 7513–7522

Figure 9: Distribution of tofu (A) and raw soybean seed (B) biomarkers plotted across all samples





Agricultural Innovation Program
Research Project Final Report

B (I). Funded Collaborators (Co-PI, AAFC, other federal scientists) • Include the name of scientist / organization.
Japanese Collaborators Mr. Mike Treytiak - KMDI International, LTD **Sevita International** John Hendrick, Export Sales Japan Mark MacDuff, Trait Development Manager Jim McCullagh, Vice President (Research) Stacey Simpkin, Research Coordinator **University of Ottawa** D. John T. Arnason - Professor Ammar Saleem – Research Associate Rui Liu – Laboratory Technician

B (II). Acknowledgement of non-funded collaborators (who provide support, e.g. access to other laboratory or other facilities and equipment input / advice / guidance / assistance, etc). • For research supported by targeted funding programs, list any collaborators who are receiving Contribution Vote 10 funds (e.g., university and industry collaborators). In addition, please list separately the participants who support your project but are not receiving any funding through the program. • Include name of scientist / organization.
Mark Berhow, USDA, Peoria, IL –soybean saponin collaborator Philippe Seguin, MacDonald College of McGill, phytochemistry collaborator

C. Variance Report (if applicable, describe how the work differs from the proposed research) • Include changes to objectives and project work plan / budget, changes to the team, other constraints.
• No changes to the objectives or project work plan.

D. Impact Assessment (if applicable, describe how the variance factors above will impact project continuation) • Include changes to the objectives, changes to the project work plan / budget, changes to performance (i.e. meeting targets).
• No changes to the objective or project work plan

E. Achievements (include only those related to this project)

Agricultural Innovation Program
Research Project Final Report

• Include innovations, publications / conferences, technology transfer, capacity building, success stories, media, recognition and other outputs.
Major achievements include: • Elucidation of 200+ new compounds in the soybean metabolome, not previously reported. • Development of methods to match metabolome of comparable varieties and distinguish distant varieties through PCA

F. Lessons learned (self-evaluation of project)
Further sensory evaluations are needed to increase the number of panel derived significant differences within taste attributes to strengthen/increase the ability of the biomarkers to be conclusive in the recommendations for suggested replacement varieties for Japanese end users.

Jim McCullagh	**May 31 2013**	
PI Name	**Date**	**Signature**

Agridocs #3783321 Final_Science_Report_-_Vote_10_Funding_.doc