

**Metabolomics Analysis  
of *Phaseolus vulgaris* L.  
Affect on a Breast  
Cancer Rat Model**

**Compound Identification Using  
an Accurate Mass Retention  
Time Library**

Steven Fischer  
Senior Application Scientist  
LC/MS Marketing  
Agilent Technologies,  
Santa Clara, CA

# Topics

Compound identification challenges

Compound identification using AMRT

- Accurate mass matching
- Retention time matching
- Isotope matching

LC/MS metabolomics example

- *Phaseolus vulgaris L.* Affect on a Breast Cancer Rat Model



# Metabolomics Compound Identification Challenges

Compounds can be classified as:

1. Known metabolite
  - Make it easy to identify
2. Known chemical but not a known metabolite
  - Make it possible to identify using MS/MS
3. Unknown chemical and unknown metabolite
  - Major project to identify



# Metabolite Identification Schemes

1. Search a mass spectral library
  - **Agilent Fiehn GC/MS Metabolomics Library**
  - Metabolomics specific LC/MS/MS Library
2. Search a database, buy the standard and re-chromatograph
  - **METLIN AMRT Database**
3. Interpret an MS/MS spectra
  - MS/MS pattern matching software
4. Purify the compound and analyze by NMR with mass spectral data support
  - Mass directed purification systems



# Compound Identification Using an Accurate Mass Retention Time Library

# Analysis Of Metabolites Using METLIN With Retention Time And Isotope Pattern Scoring

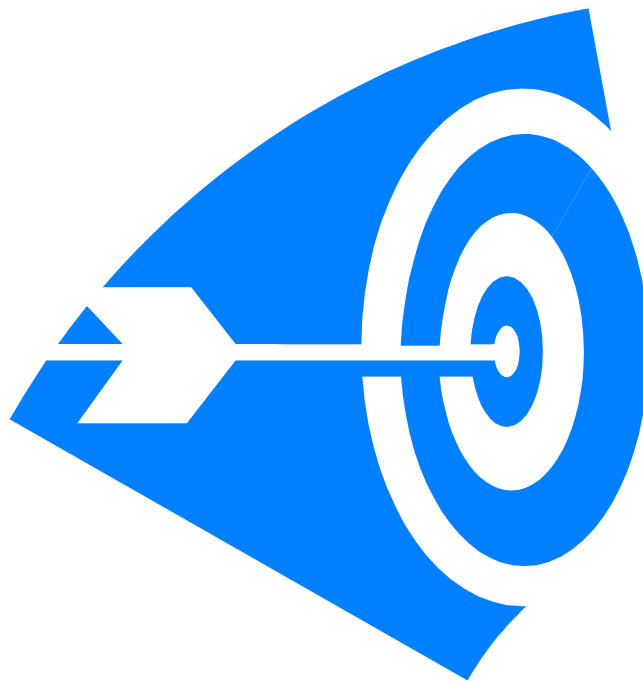
Increase specificity of identification

- Search on mono-isotopic accurate mass
- Require retention time
- Calculate empirical formula using full isotope pattern

Generate empirical formula for unmatched data

- Add found but unknown compound to database for future tracking
  - retention time and proposed formula

Molecular Formula and METLIN Personal Metabolite Database Matching Applied to the Identification of Compounds Generated by LC/TOF-MS, T. Sana et al, *Journal of Biomolecular Techniques* 19:258–266 2008



# METLIN Personal Database Overview

Metabolite-specific database for metabolomics research

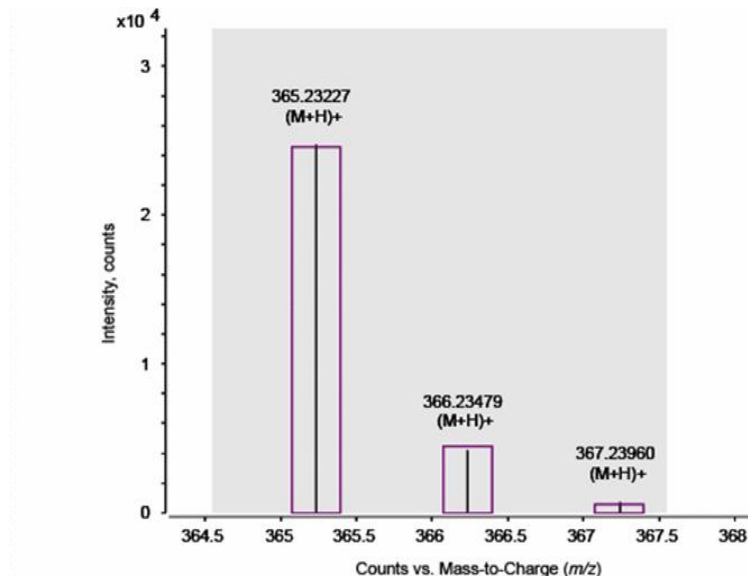
- Database installed on a PC
- Contains ~22000 compounds
- Manual and batch searches
  - Query based on monoisotopic mass
- Customizable
  - Add compounds
  - Assign chromatographic retention times to metabolites
  - Create subset databases
- Works with other Agilent software



# Calculating a Molecular Formula (MFG) With Database Searching

## Dihydrocortisol example

- Database search
  - Uses only mono-isotopic mass; loses isotope information
  - Only compounds in the database can be matched; database selection creates specificity based on biology
- Molecular formula calculation (MFG)
  - Uses mass values of all isotopes, including adducts
  - Support database match; if no retention time information present
  - Provides useful information; even if no database match



MS Formula Results: Scan [ 9.683-19.735 min ]

Peak	Formula [v]	Ion Formula	Score	Mass	Calc. Mass	Difference (amu)	Abt. Diff (amu)	DFE
1	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	100	364.225	364.22334	-2.09	2.00	
2	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	95.7	364.225	364.22535	0.78	3.79	
3	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	85.5	364.225	364.22630	5.32	5.32	
4	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	76.42	364.225	364.22437	-0.04	3.08	

Isotope	Abund%	Calc. Abund%	m/z	Calc. m/z	Difference (amu)
1	100	100	365.23227	365.23227	0.00
2	17.05	22.26	366.23479	366.23564	2.89
3	2.83	3.62	367.23960	367.23834	-3.42

Peak	Formula [v]	Ion Formula	Score	Mass	Calc. Mass	Difference (amu)	Abt. Diff (amu)	DFE
1	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	71.52	364.225	364.22630	1.92	1.92	
2	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	61.32	364.225	364.22228	-7.43	7.43	
3	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	56.77	364.225	364.22329	-5.22	5.22	
4	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	55.78	364.225	364.22631	3.6	3.61	
5	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub>	50.54	364.225	364.22737	-5.55	5.55	



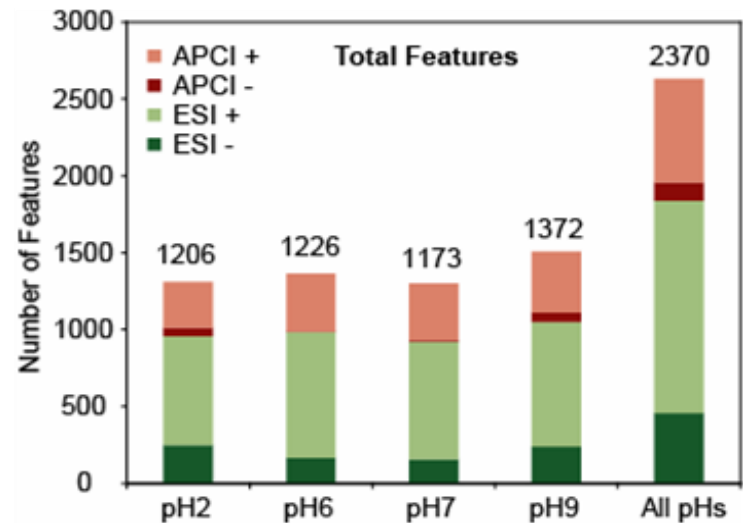
# Erythrocyte Metabolites Extracted At Different pH And Ionization At Different Modes

No single ionization mode detects all analytes

1. ESI 67% vs. APCI 25%
  - Common to both 7.3%
2. Pos mode 74% vs. Neg mode 22%
  - Common to both 3.9%

Need method for ESI (+/-) and APCI (+/-)

A sample extraction and chromatographic strategy for increasing LC/MS detection coverage of the erythrocyte metabolome, T.R. Sana et al., *J. Chromatography B* 871 (2008) 314-321



Chromatography method must be compatible with ESI & APCI and positive & negative ionization

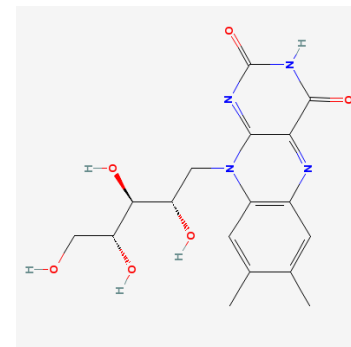
# AMRT Analytical Methodology

- Single reverse phase separation method
  - Water, methanol and acetic acid
  - Reverse phase column
    - 2.1 x 50 mm, 1.8  $\mu\text{m}$  Zorbax SB-Aq
    - 2.1 x 30 mm, 3.5  $\mu\text{m}$  Zorbax SB-C8
  - Flow rate 0.6 ml/min
  - Simple linear gradient
    - 2% methanol to 98% methanol in 13 minutes
    - 6 minute hold at 98% methanol
  - Cycle time 24 minutes
- Compatible with ESI / APCI and positive / negative ionization modes

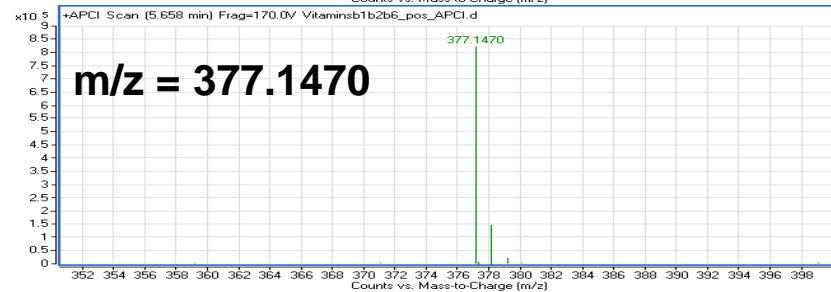
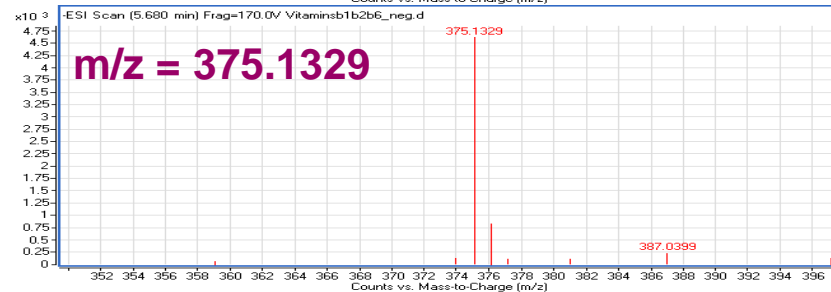
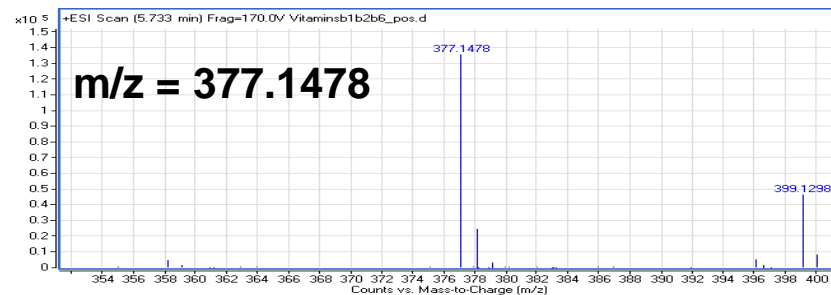
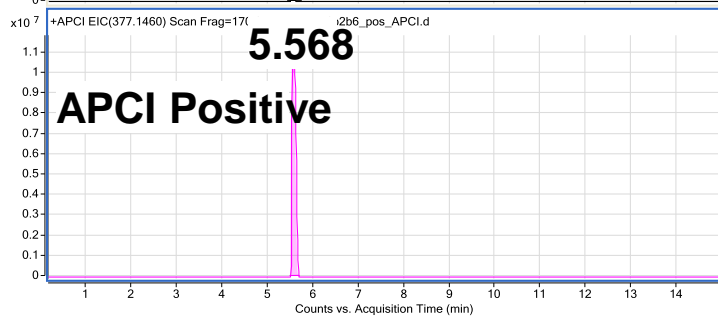
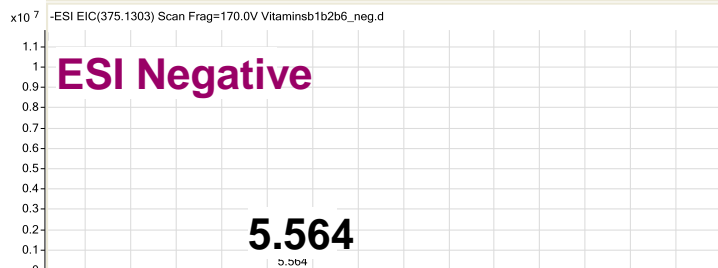
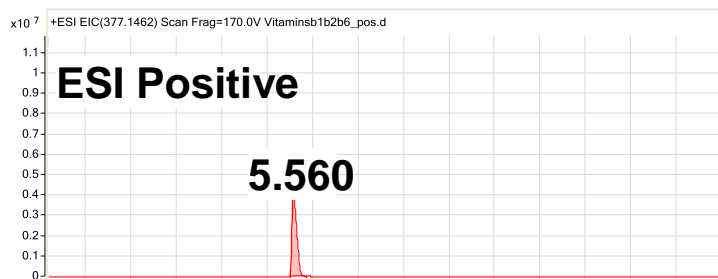


# Vitamin B2 Standard: $C_{17}H_{20}N_4O_6$

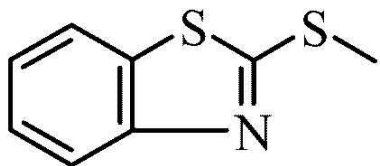
$[M+H] = 377.1456$ ,  $[M-H] = 375.1310$



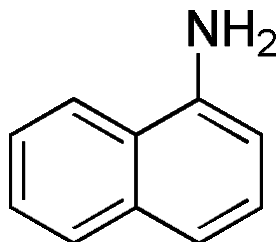
$RT_{ave} = 5.564$



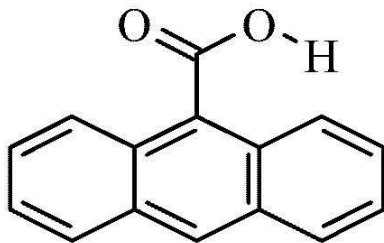
# LC/MS Internal Standards



2-(methylthio)benzothiazole



1-naphthylamine



9-anthracene carboxylic acid

## 2-(methylthio)benzothiazole

- APCI only – positive ion only
- Recovery has significant dependence on pH and salt

## 1-naphthylamine

- APCI and ESI – positive ion only
- Good solubility, especially at low pH

## 9-anthracene carboxylic acid

- APCI and ESI – positive and negative ion mode
- Good solubility, especially at high pH

# Using Accurate Mass, Retention Time Database Searching With Molecular Formula (MFG)

Human Urine Sample 11

Open... Save Save As... Export Close

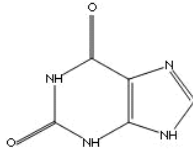
Batch Result Batch Summary

1070 Features

	Mass	RT	# DB Hits	# MFG Candid.	Feature ID	Abundance	log2(I)
485	219.1100	4.019	0	0	2	162	273556
486	99.0689	4.092	0	0	1	4	9118307
▶ 487	152.0333	4.146	1	1	3	196	213279
488	168.0536	4.149	0	0	1	712	24296
489	111.0684	4.196	0	0	1	950	12353
490	174.1001	4.196	0	0	1	729	22760
491	216.1101	4.198	0	0	2	246	159829
492	99.0686	4.261	0	0	1	34	1788792

Structure Information

Structure MOL Text



Notes

Endogenous Metabolite

Geigy vol. 3 p. 103

Mass = 152.0333, Time = 4.146

	Best	Obs. Mass	Obs. RT	Name	Formula	ΔMass(ppm)	ΔRT	MFG Score	CAS	METLIN	KEGG	HMP
▶ 1	<input checked="" type="checkbox"/>	152.0334	0.000	Xanthine	C5H4N4O2	0.4	4.130	100.0	69-89-6	82	C00385	
2	<input type="checkbox"/>	152.0334	0.000		C5H12O52	-2.7		91.7				
3	<input type="checkbox"/>	152.0334	0.000		C4H8O6	-8.4		59.3				
* 4	<input type="checkbox"/>											

Results are returned prioritized by;

- AMRT match and MFG
- AMRT only
- Database match and MFG
- Database match
- MFG

# Reverse Phase, Accurate Mass, Retention Time (AMRT) Metabolite Database

The screenshot displays the Agilent METLIN Personal Metabolite Database interface. The window title is "Agilent METLIN Personal Metabolite Database - C:\Documents and Settings\Wischers\Desktop\Chris Metabolomics Data\Metlin-Urine.mtl". The interface includes a menu bar (File, Database, Edit Metabolites, View, Metabolomics Links, Help) and a toolbar with icons for finding and editing metabolites. The main area is divided into several sections:

- Search and Edit Fields:** Name: Hippuric acid; Mass: 173.05824; RT: 9.409; Formula: C9H9NO3; CAS: 495-69-2; METLIN: 1301; KEGG: C01586; HMP: ; Metabolite ID: 1267.
- Radical ion type:** Radio buttons for Neutral (selected), Anion, and Cation.
- Edit actions:** Buttons for Add New, Save As New, Update Selected, and Delete Selected.
- Molecule:** Structure and MOL Text tabs. The Structure tab shows a chemical structure of Hippuric acid.
- Notes:** Antidepressant, MAO inhibitor; Metabolite of Tranylcypromine and Benzyl benzoate; Dolly, Colin Therapeutic Drugs, 2nd Ed. 1999 p. B34, T154.

Below the main area, a "Single Search Results: 1 hit for" section contains a table with the following data:

Name	Formula	Mass	Cation	RT (min)	CAS	METLIN	KEGG	HMP
Hippuric acid	C9H9NO3	173.05824	<input type="checkbox"/>	9.409	495-69-2	1301	C01586	

Retention times were added to the database by running standards

- Database is separation specific; retention times change if separation is changed
- Compounds need to be retained ( $k' > 2$ ) in order for retention time to add specificity
- Compounds of same mass ranked differently if retention time information present
- RP-AMRT currently contains 363 database entries with retention time

**Metabolomics Analysis  
of *Phaseolus vulgaris* L.  
Affect on a Breast  
Cancer Rat Model**

Steven Fischer  
LC/MS Marketing  
Agilent Technologies,  
Santa Clara, CA

Henry J. Thompson  
Meghan M. Caulum  
Cancer Prevention Laboratory  
Colorado State University

# Dry Bean / Rat Tumor Study

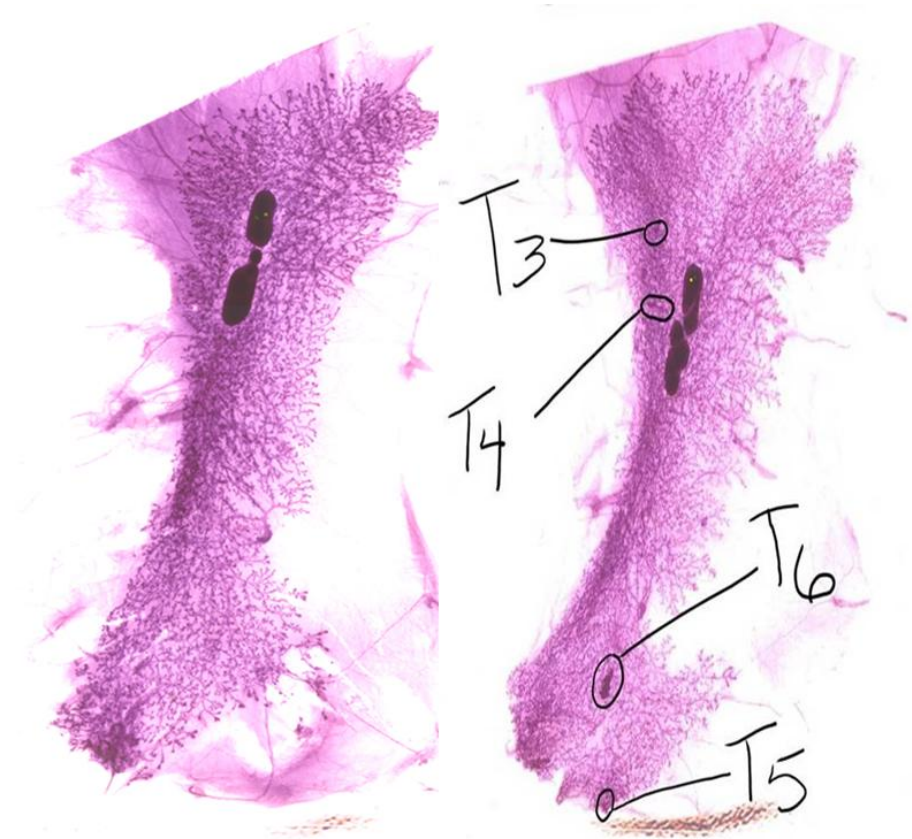
Breast cancer rat model responds to different dry beans

- White kidney bean is most effective
- Navy bean is intermediately effective
- Red bean is partially effective
- Rat chow has no effect

Are there observable chemical differences in beans?

Can the differences be seen in plasma, gland or tumor?

Treatment	Sample Number	incidence (%)	average tumors per	Tumor burden g/rat
Control	30	93.3	3.2	0.48
Navy Bean	30	70.0	1.9	0.34
White Kidney Bean	30	66.7	1.0	0.05



**Tumor free  
mammary gland**

**Tumor infiltrated  
mammary gland**



# Study Design

## Diet test

- 10 lots produced of each
  - House blend rat chow AIN-93G
  - 60 % white kidney to chow blend
  - 60 % navy to chow blend
  - 60 % red to chow blend

## Animal - Diet test

- 10 animals feed diet of
  - House blend rat chow AIN-93G
  - 60 % white kidney added to chow
  - 60 % navy added to chow
  - 60 % red added to chow



# Sample Processing Protocol

## Sample extraction

- Diet 50 mg, Plasma 100 ul, tissue 100 mg
- water / methanol / chloroform
  - pH 2 and pH 9

## Extract preparation

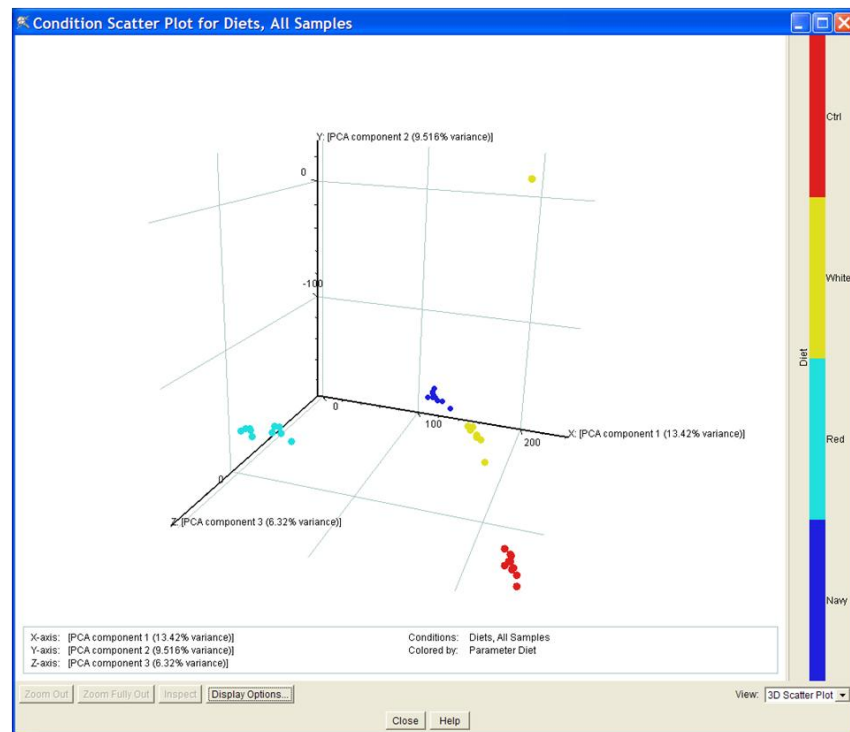
- Speed-Vac to dryness
- Protein precipitate with 75% ACN
- Speed-Vac to dryness
- Reconstitute with 50 ul methanol, vortex then 50 ul water, vortex
- Blend pH 2 and pH 9 extracts
- Inject 5 ul into LC/MS system



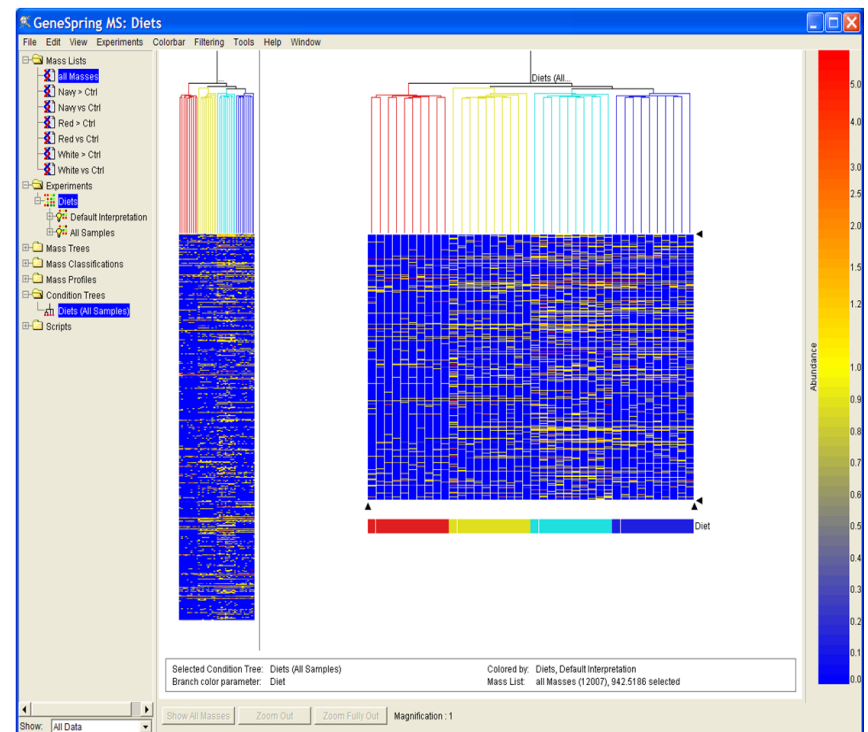


# Initial Analysis Of Bean Supplemented Diets Data

## PCA of Bean Diet Data



## Condition Tree of Bean Diet Data



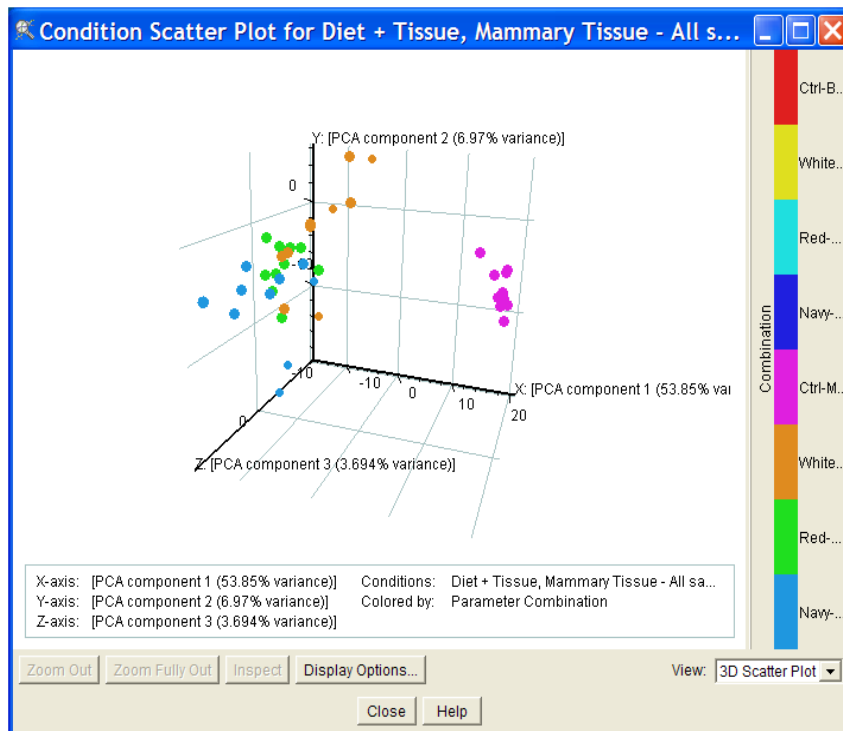
# Identification Results Of Bean Diet Samples

## LC/MS; ESI-Positive Ion

	Control		White		Navy		Red	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total Features	692	100.0%	1386	100.0%	1639	100.0%	2290	100.0%
DB+RT+MFG	7	1.0%	10	0.7%	10	0.6%	14	0.6%
DB+RT	0	0.0%	2	0.1%	1	0.1%	5	0.2%
DB+MFG	89	12.9%	322	23.2%	340	20.7%	481	21.0%
DB	31	4.5%	69	5.0%	90	5.5%	144	6.3%
MFG	444	64.2%	655	47.3%	716	43.7%	1025	44.8%
No match	121	17.5%	328	23.7%	482	29.4%	621	27.1%
	692		1386		1639		2290	

# Analysis Of Rat Mammary Gland After Four Different Bean Diets

## PCA of Mammary Gland Data



## Condition Tree of Mammary Gland Data



# Identification Results Of Mammary-Bean Diet Samples, LC/MS; ESI-Positive Ion

	Control		White		Navy		Red	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total Features	1218	100.0%	1186	100.0%	1182	100.0%	1141	100.0%
DB+RT+MFG	10	0.8%	10	0.8%	10	0.8%	8	0.7%
DB+RT	0	0.0%	1	0.1%	1	0.1%	1	0.1%
DB+MFG	287	23.6%	284	23.9%	319	27.0%	272	23.8%
DB	18	1.5%	8	0.7%	8	0.7%	27	2.4%
MFG	743	61.0%	707	59.6%	688	58.2%	693	60.7%
No match	160	13.1%	176	14.8%	156	13.2%	140	12.3%
	1218		1186		1182		1141	

# Bean Metabolites Found In Mammary Gland Tissue

## White Bean

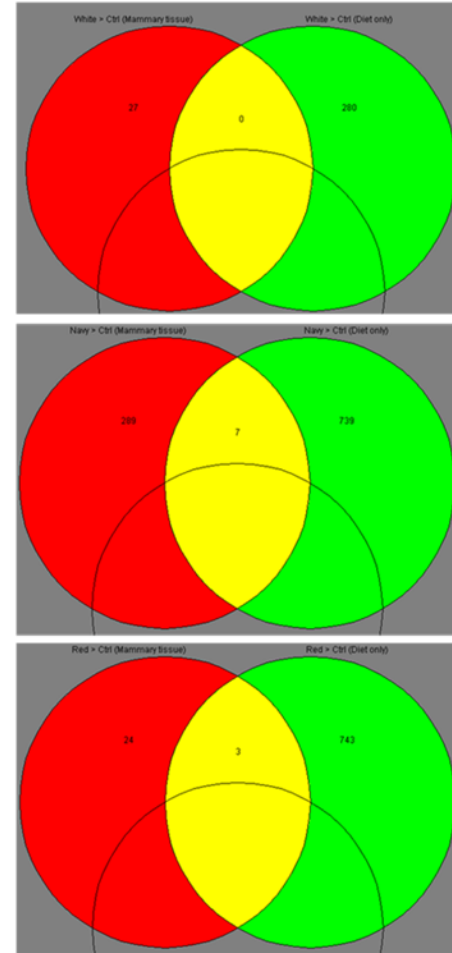
- Present in both – **0**
- Unique to tissue – **27**
- Unique to bean – **280**

## Navy Bean

- Present in both – **7**
- Unique to tissue – **289**
- Unique to bean – **739**

## Red Bean

- Present in both – **3**
- Unique to tissue – **24**
- Unique to bean – **743**





# AMRT Search Results For Common Metabolites

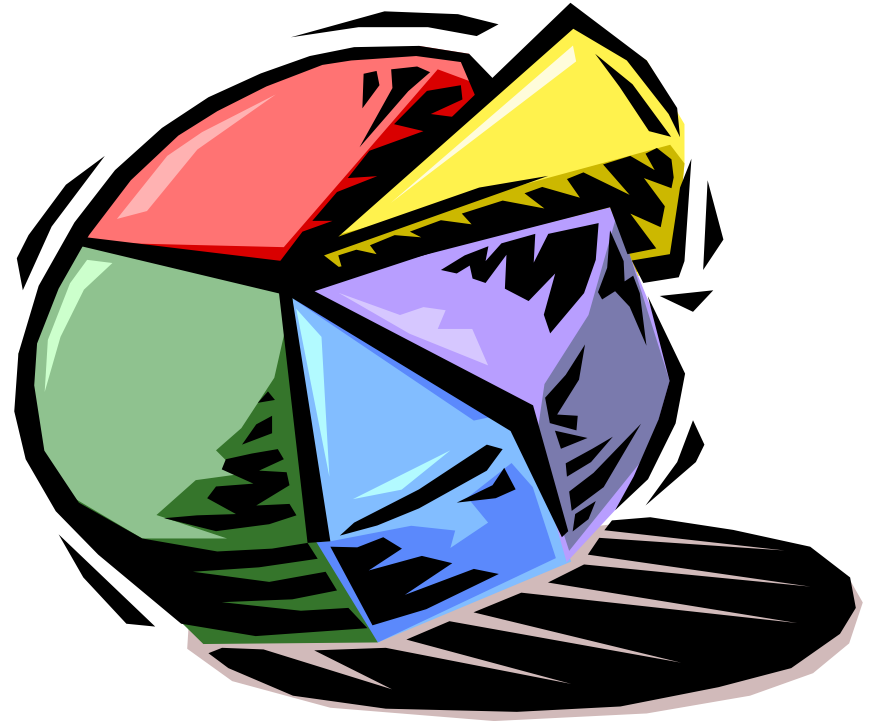
White Bean – 0

Navy Bean – 7

- DB + RT + MFG – 1 (niacinamide)
- DB + MFG – 2
- MFG – 4

Red Bean – 3

- DB + RT + MFG – 0
- DB + MFG – 2
- MFG - 1



# Bean-Rat Model Observations With Regard to Identification

- 1 in 4 features were found in the METLIN database
  - A RT database of 363 compounds is too small (~4% of DB matches)
- 4 in 7 features had only an empirical formula calculated
- 1 in 7 features were complete unknowns



