



Improving Lead Discovery from Natural Products

D. J. Livingstone*, F. Giaquinto and N. Porter

*ChemQuest, Isle of Wight, UK
& BioDiversity Ltd., London, UK



Combinatorial chemistry

An internet query -

‘What is combinatorial chemistry?’

and a reply -

‘A method of increasing the size of the haystack in which you find a needle.’ *

*C. D. Floyd *et al.* (1996). *Chemistry in Britain*, March, pp.31-5.



Drugs from Natural Products

★ Many naturally derived drugs have been in use for thousands of years

★ Well known examples are:

- ◆ Morphine (~4000 B.C.)
- ◆ Reserpine (<1000 B.C.)
- ◆ Aspirin (<200 B.C.)
- ◆ Ephedrine (<1 A.D.)
- ◆ Quinine (<1650)
- ◆ Digitalis (1775)

Are natural products still important ?

★ In a recent (17/2/99) BBC radio programme “Murder, Magic and Medicine” Melanie O’Neill from the dept. of Compound Diversity at GlaxoWellcome said:

‘Over 50% of prescription drugs owe their origin to plants’

Recent Drugs from Natural Products

- ★ Cyclosporin
- ★ Mevinolin (HMG CoA Reductase)
- ★ Captopril, Enalapril (snake venom)
- ★ Taxol (winner R&D 100 award 1995)
- ★ Integrilin (anticoagulation - approved 1998)

Natural Product Sources

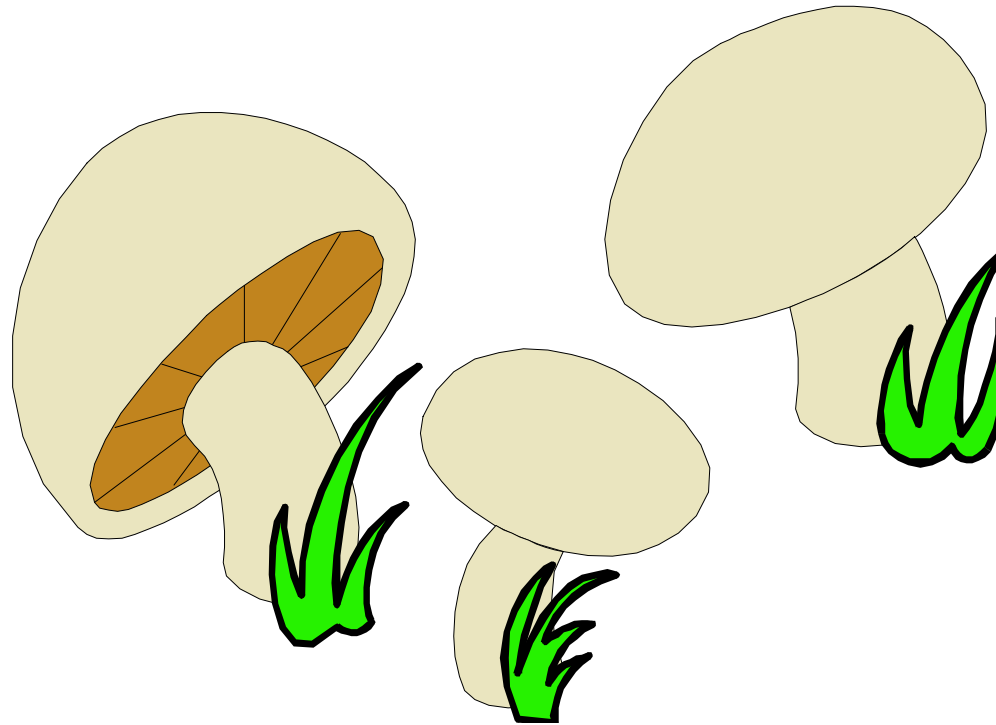
- ★ Extracts from organs and tissues from animals, insects, molluscs, *etc.*
- ★ Extracts from plants, seeds, fruits, *etc.*
- ★ Extracts from bacterial and fungal cultures.
- ★ Control of culture conditions is essential to the successful generation of diverse chemical structures found in secondary metabolites.



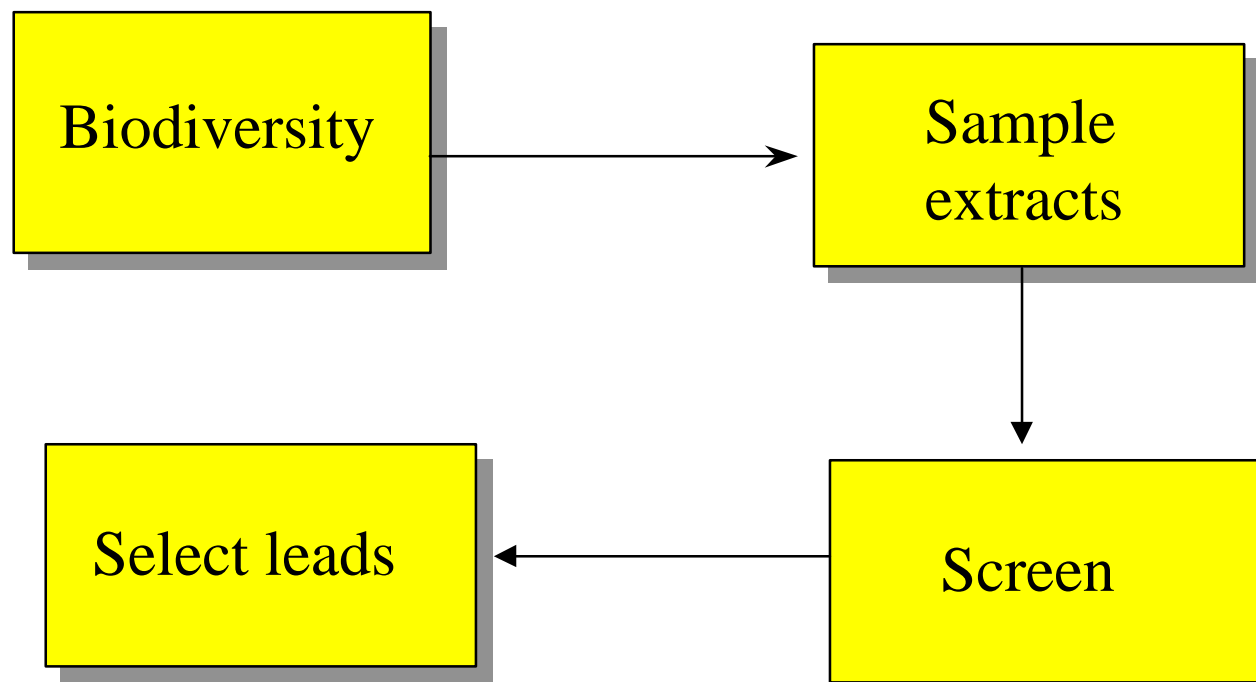
Fungi

- ★ There are 1.5 million species
- ★ Less than 1% of these are in culture
- ★ Since they are more complex than bacteria they represent a greater source of diversity

***The original talk showed some
slides of Fungi***



Natural product screening



Biodiversity as a source of leads

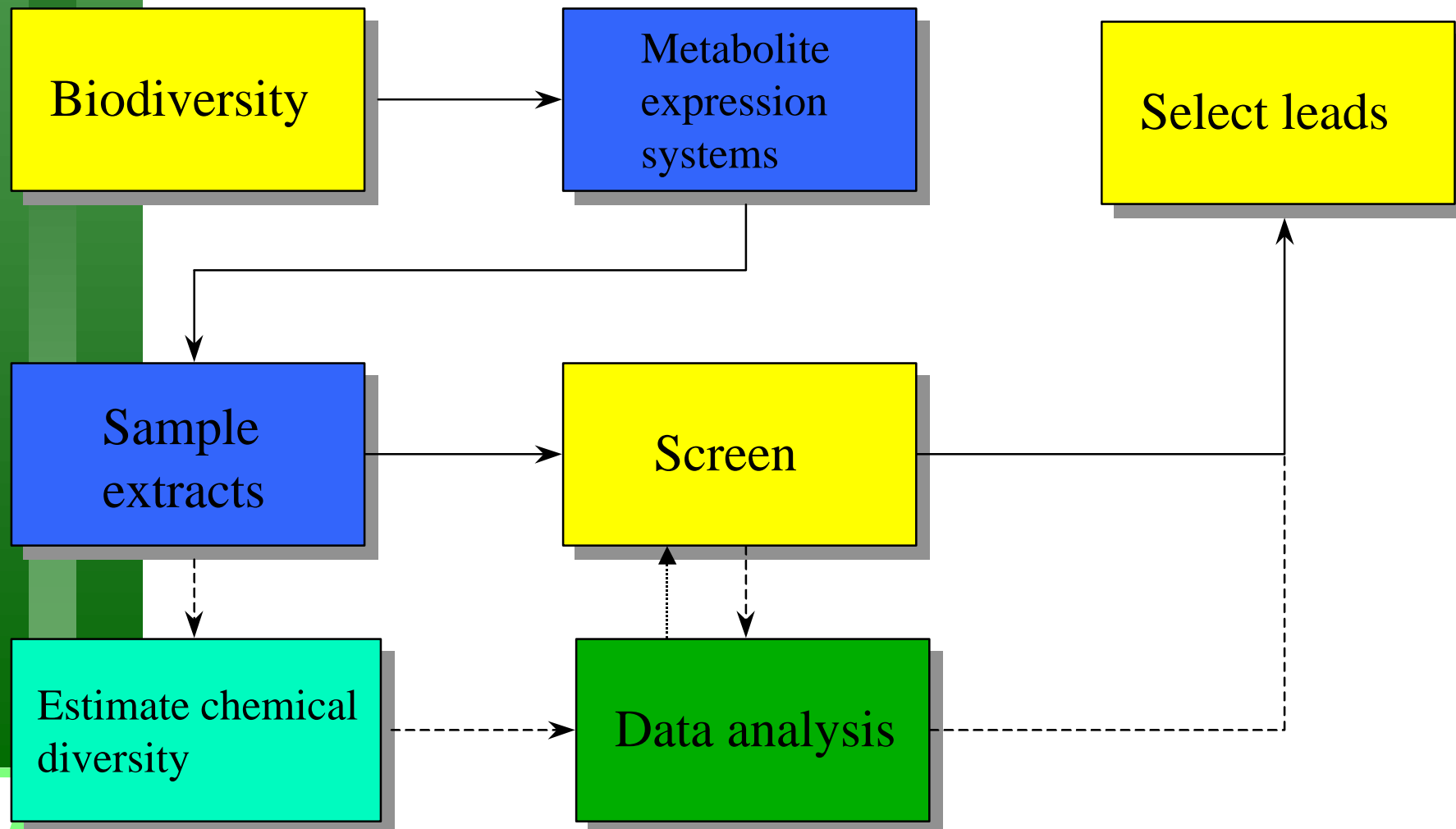


Disadvantages;

- ◆ The extracts contain many different compounds
- ◆ Chemicals range from low to very high molecular weights
- ◆ Characterisation of the extracts is difficult without separation (costly and time consuming)

Natural product screening

Improving lead selection





Estimating chemical diversity

Chemical fingerprinting

An ideal analytical system should be:

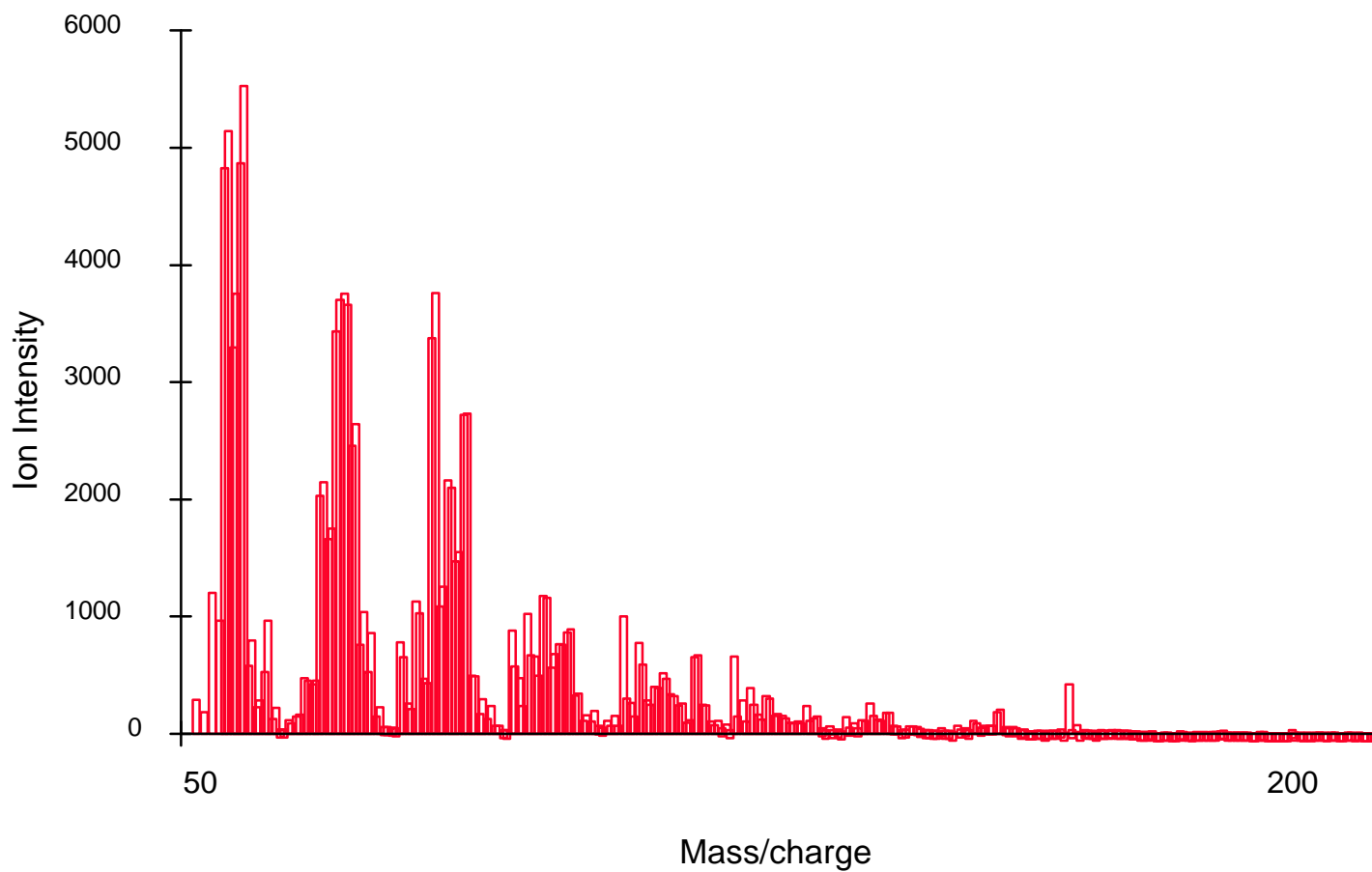
- ★ Reproducible
- ★ Simple to use
- ★ Rapid
- ★ Sensitive
- ★ Amenable to automation

Estimating chemical diversity

Chemical fingerprinting techniques

- ★ Fourier transform infrared spectrometry (FT-IR)
- ★ Ultraviolet-resonance Raman spectroscopy (UVRRS)
- ★ *Mass spectrometry (MS)*
 - ◆ PyMS
 - ◆ *MALDI-TOF MS*

Pyrolysis mass spectrum



PyMS Properties

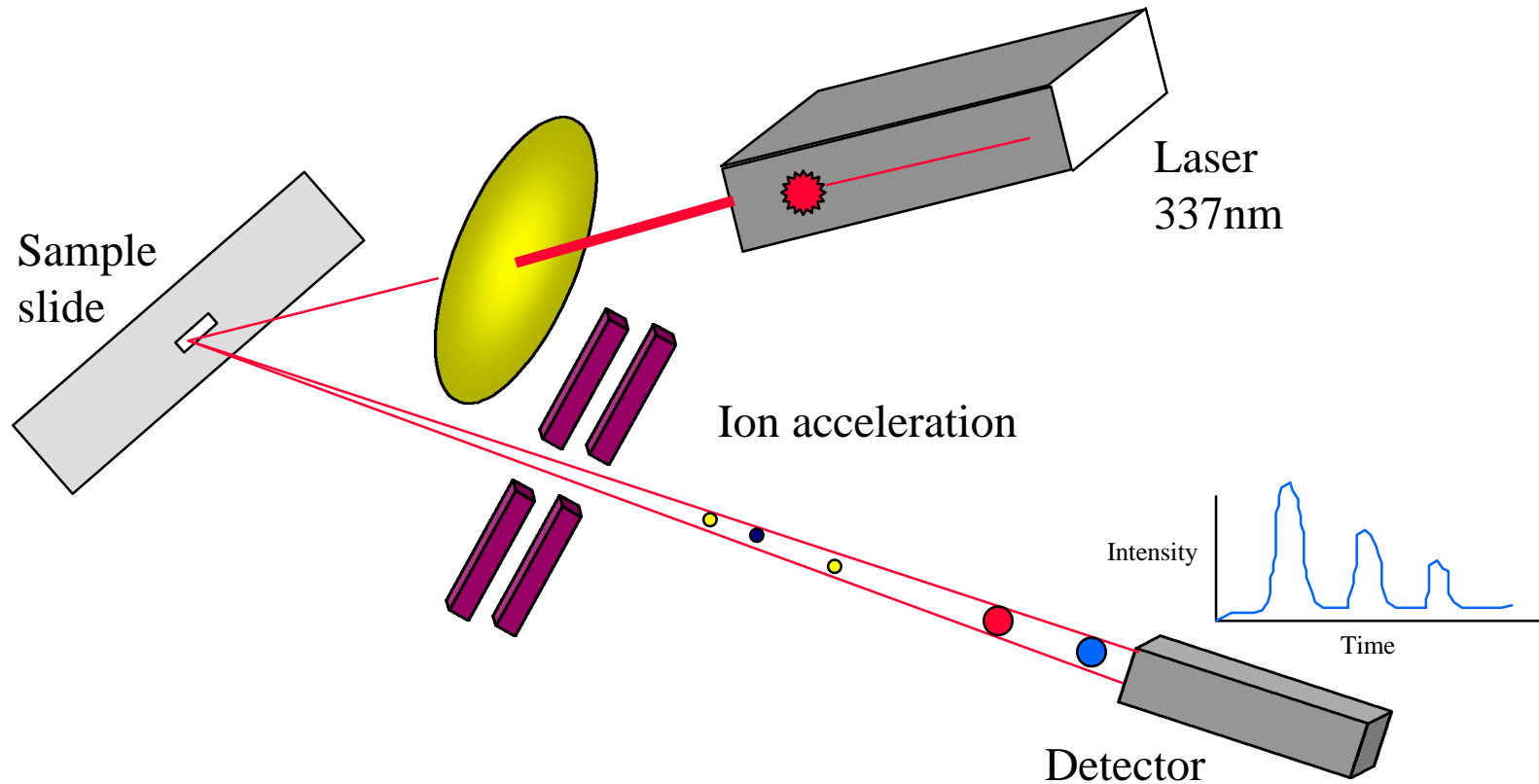
- ★ Limited mass range: measurements typically over 50-200 Da
- ★ Spectra consist of superimposed fragmentation patterns of individual components
- ★ Information lies in the mass patterns
- ★ Subtle differences not discernible by eye but detected by mathematical analysis

Major Drawbacks

- ★ All information held in the overall mass fragmentation pattern - no molecular ions
- ★ Sample preparation cumbersome
- ★ Poor reproducibility (day to day, week to week drift) does not allow the establishment of data libraries

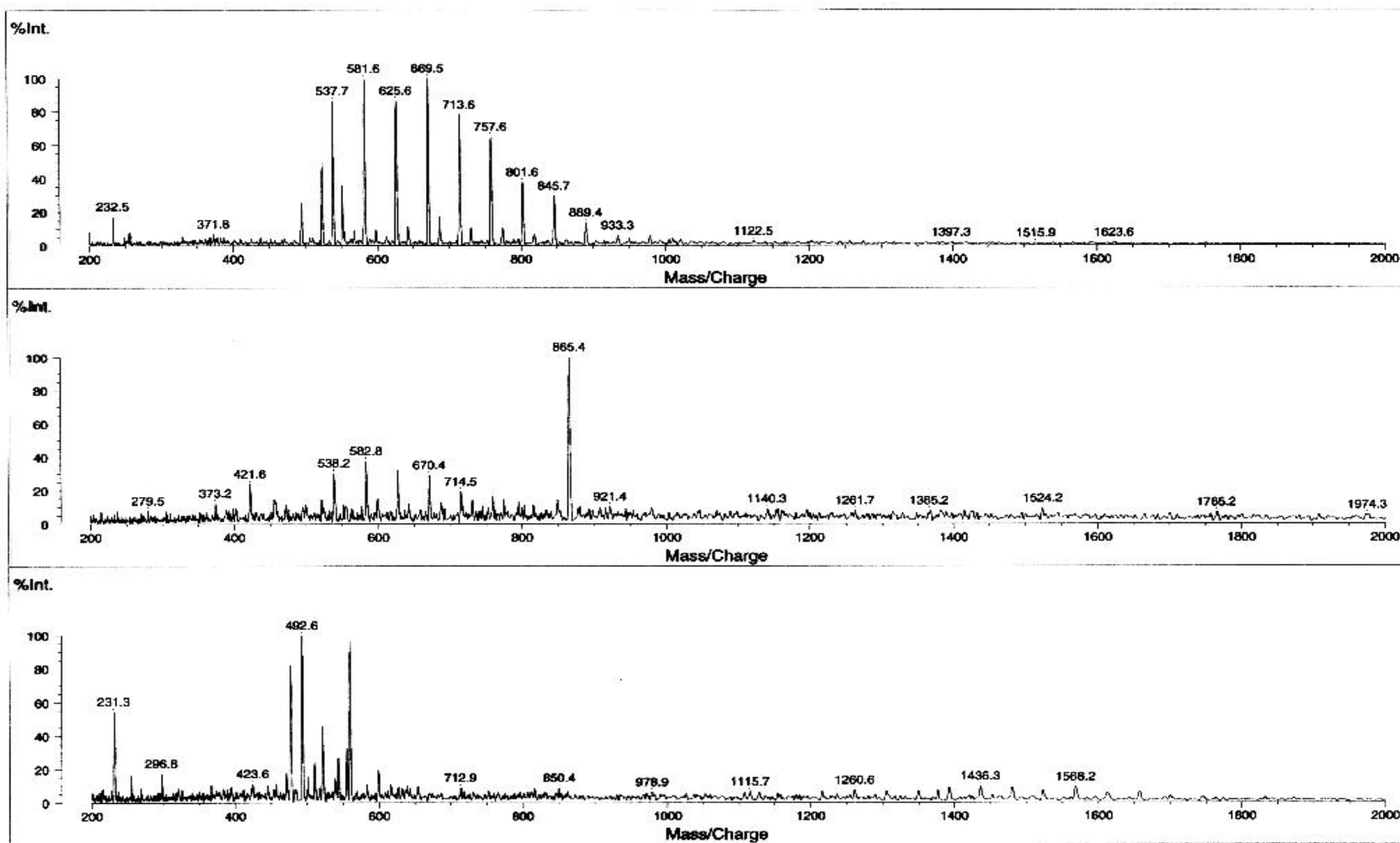
MALDI-TOF MS

Schematic



Time of flight achieves both separation and measurement of all components initially ionized by the laser/matrix combination

MALDI-TOF Spectra



MALDI-TOF MS

Matrix assisted laser desorption ionisation time of flight mass spectrometry

- ★ Applicable to complex mixtures
- ★ Mass range up to 250,000 Daltons
- ★ Can generate and measure ions of intact molecules
- ★ Data is highly discriminative
- ★ Technology still developing
 - ◆ Matrix development (sample presentation)
 - ◆ Data analysis
 - ◆ Automation

Making Use of Spectral data

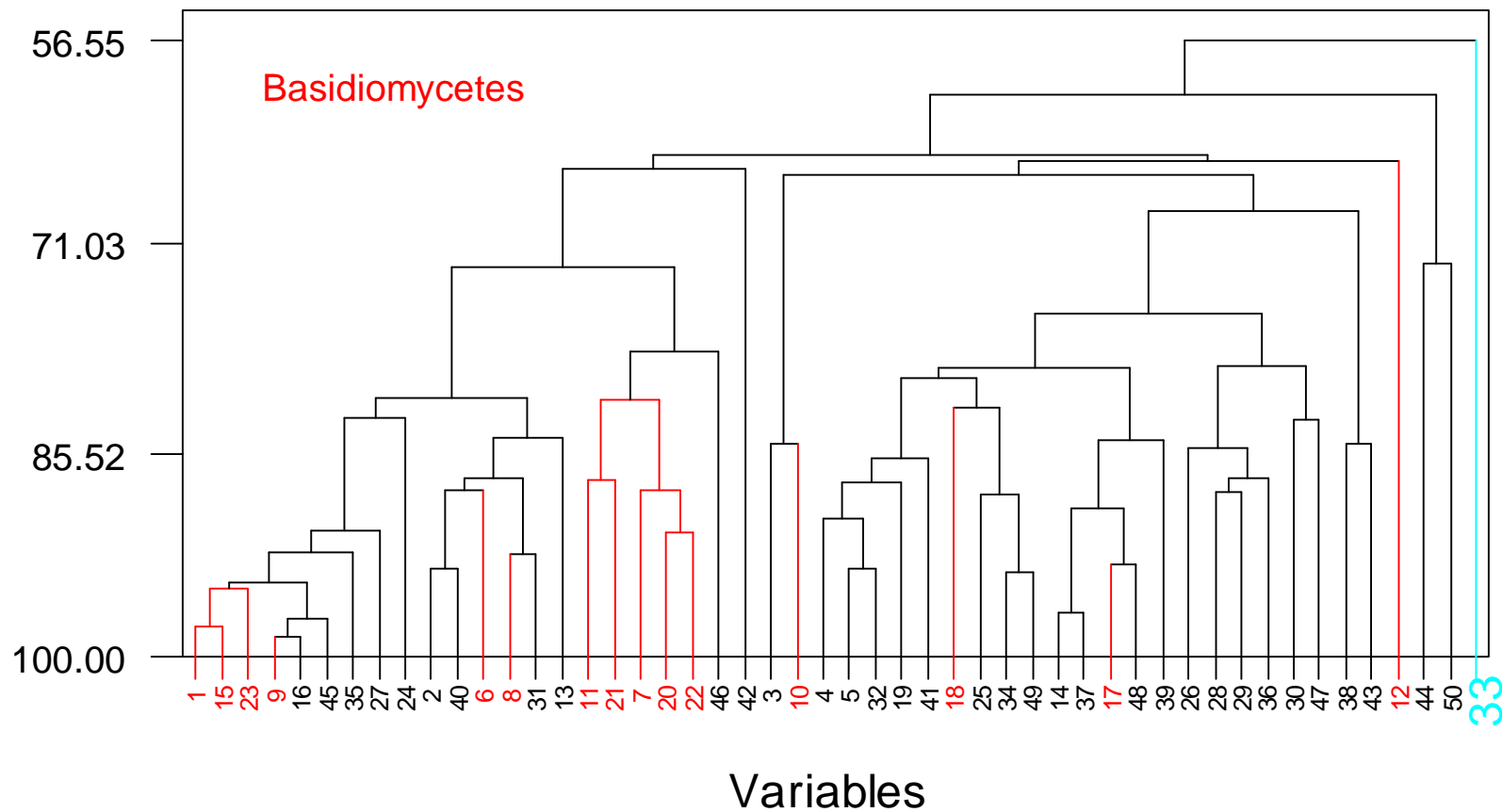
- ★ The spectral information is multivariate (thousands of m/z values).
- ★ We need to remove redundancy and “noise” from the data (preprocessing)
- ★ Mathematical and/or statistical methods are needed to “unravel” this information.
- ★ There is no “best” method to use.

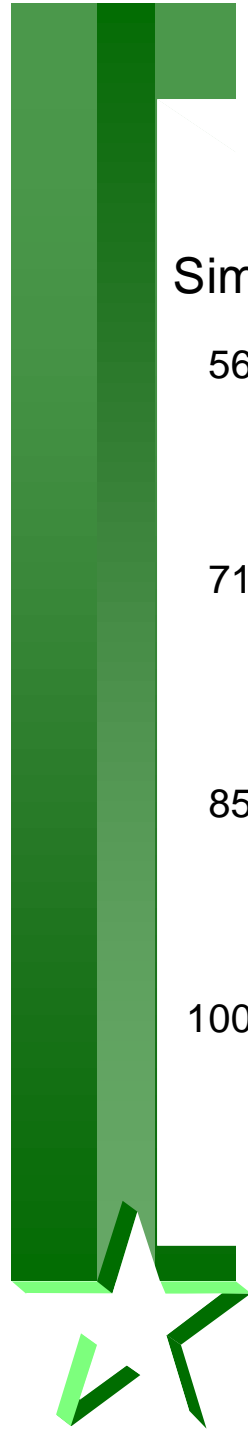
Cluster Analysis

- ★ 50 Samples from 4 classes of fungi plus 1 unknown and 1 sample from a single class.
- ★ The samples are described by MALDI-TOF data over the 200 to 2000 m/z range using 15000 values.
- ★ Cluster analysis carried out on the unstandardised raw data using average linkage (avge distance between clusters)

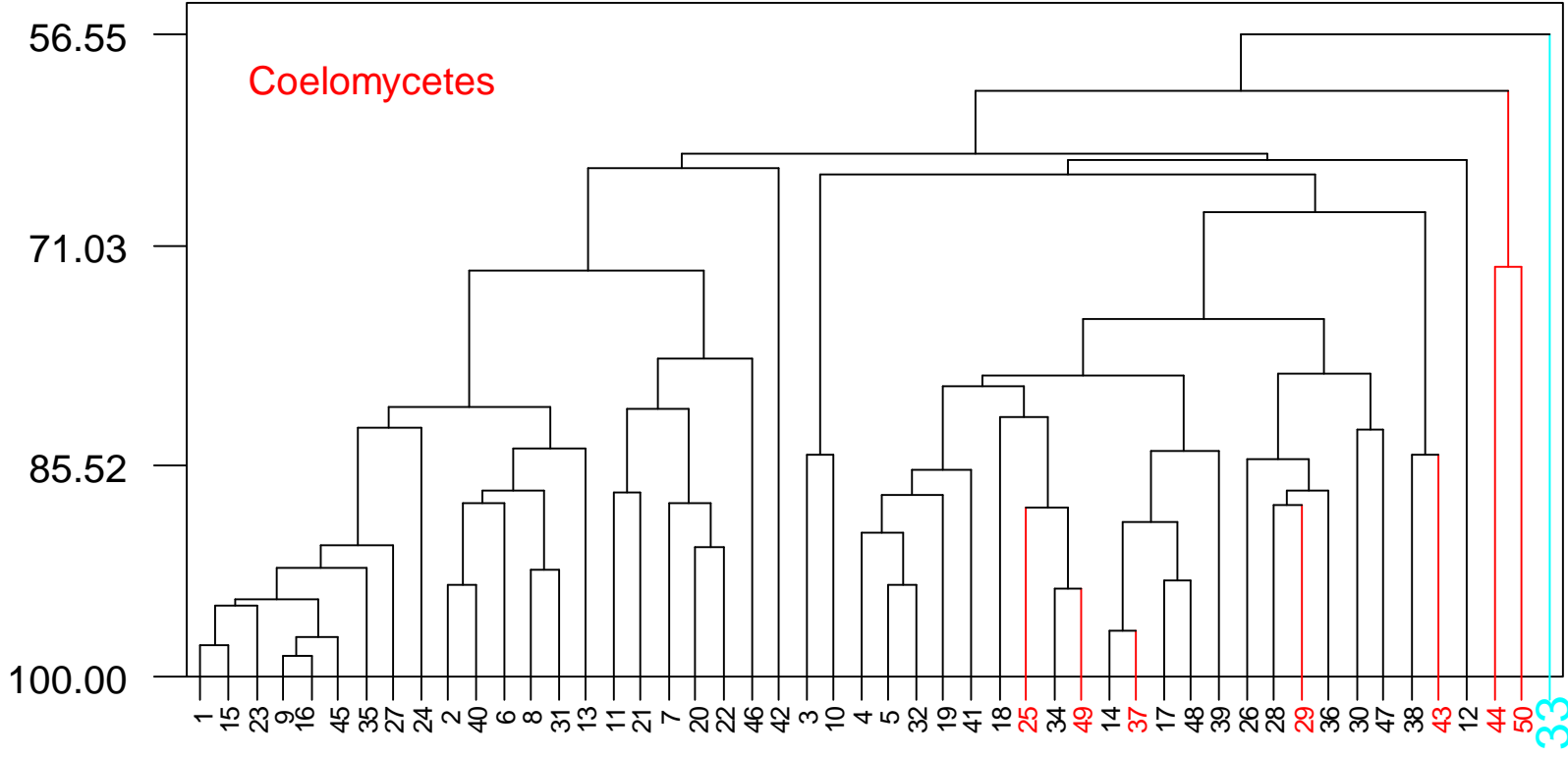
CO extracts

Similarity

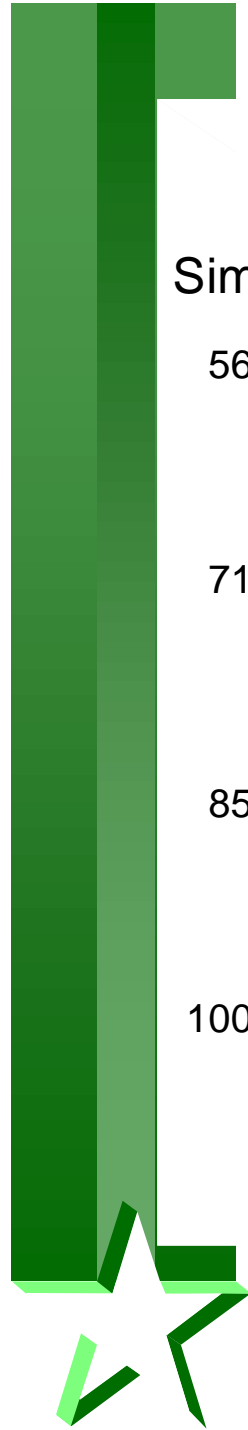




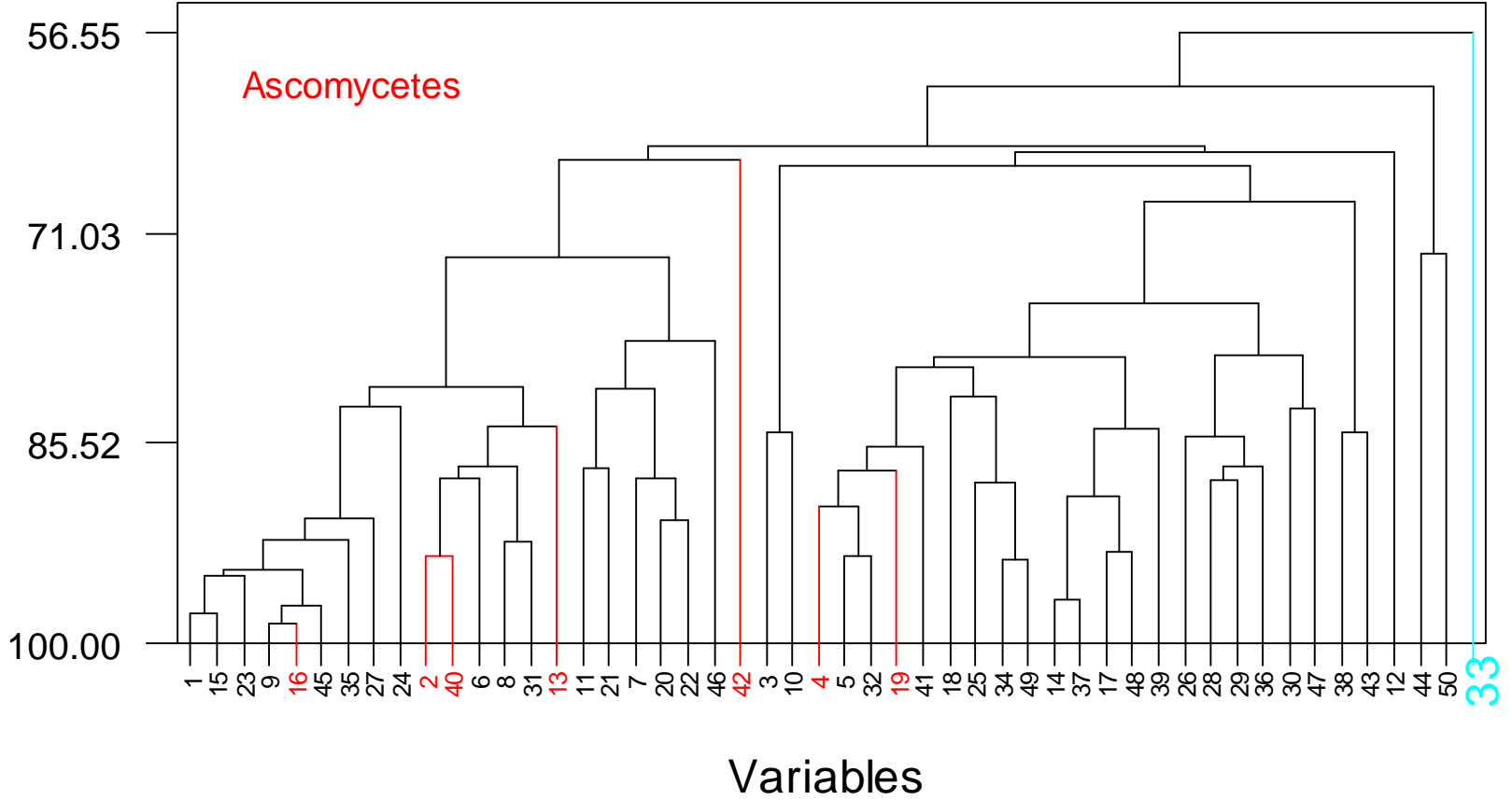
Similarity



ChemQuest




Similarity



Preprocessing

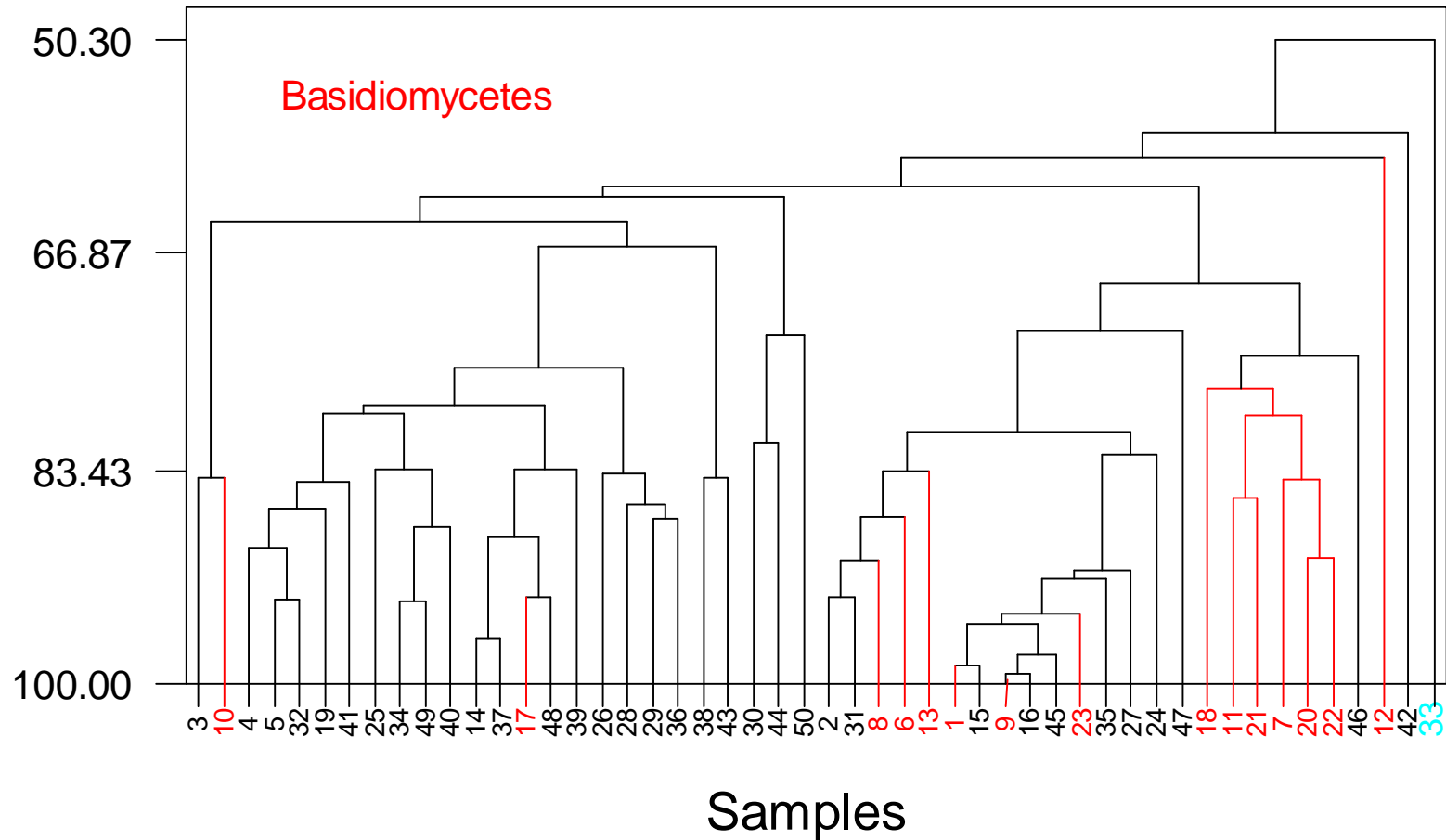
- ★ Each spectrum consists of a voltage (from the detector) over time. This equates to the amount of material at each M/Z value.
- ★ The spectra were processed by removing peaks which were less than 1, 2 and 5% of the maximum peak.
- ★ This removed ~3,000, 5000 and 9000 data points

Comparison of sets

- 
- Principal component analysis on the raw and reduced sets gave very similar results:
- ◆ 9 Components with Eigenvalues greater than 1
 - ◆ Total variance explained ranges from 85.4 % (raw) to 83.8% in the first 10 components
 - ◆ Scree plots show a similar discontinuity at 3, 6 and 8 components


5% Removed

Similarity



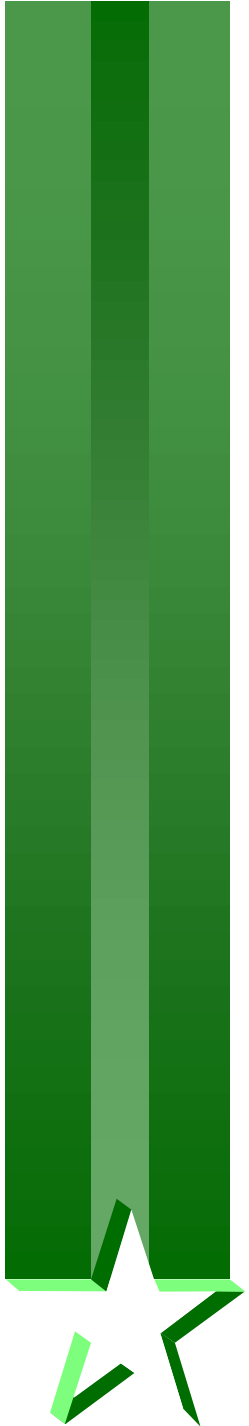
ChemQuest

Matrices

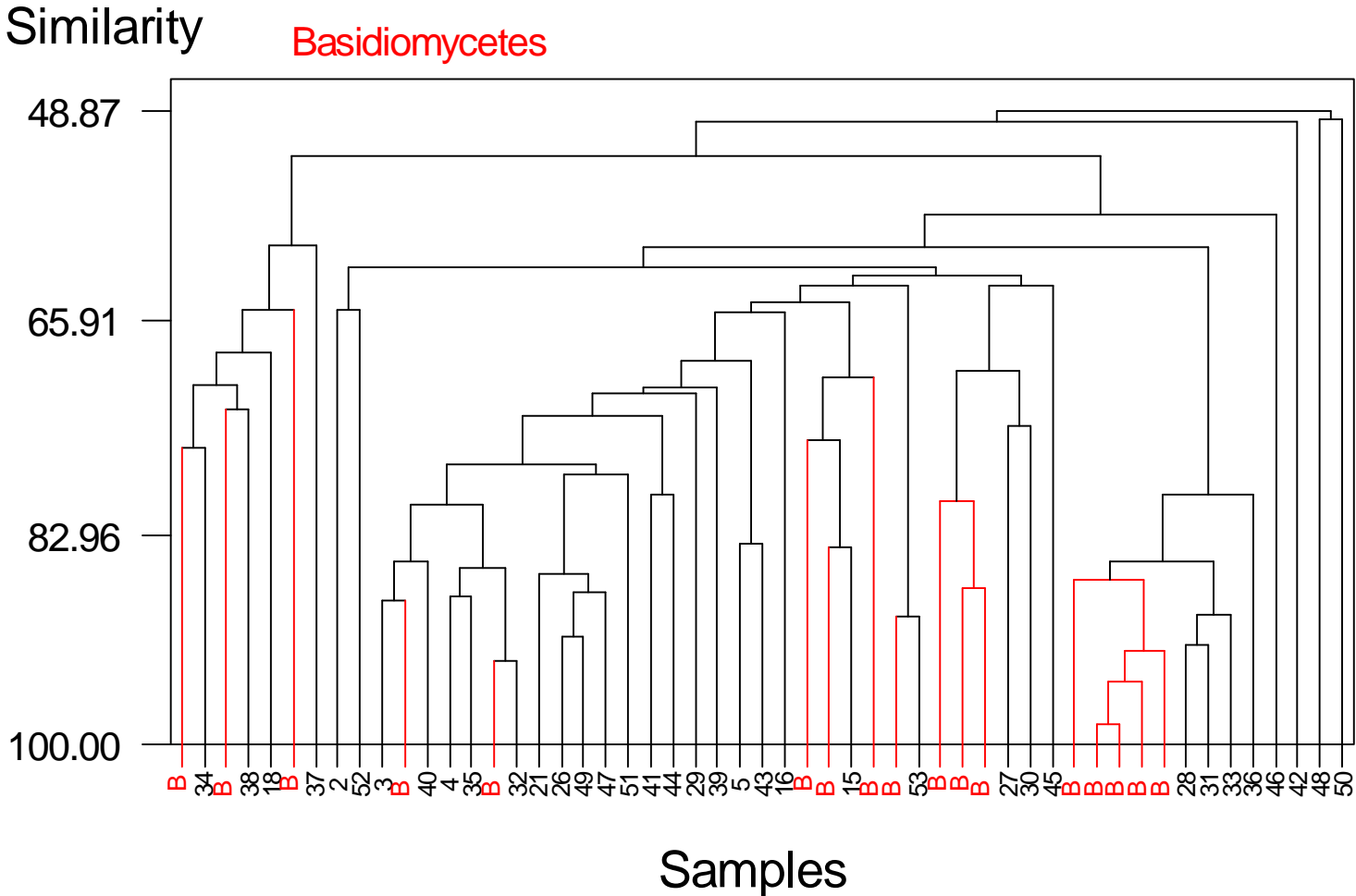
- 
- The nature of the matrix used for sampling is of great importance because it determines:
- ◆ Uniformity of distribution and thus reproducibility
 - ◆ Efficiency of the transfer of laser energy
 - too high gives sample breakdown
 - too low gives poor ionisation/flight
 - ◆ Ionisation, since the matrix takes part in reactions

Matrices

- ★ 2 matrices have been examined so far
- ★ The second matrix appears to contain more peak information:
 - ◆ Application of the 1, 2 and 5% filters resulted in the removal of no data
 - ◆ Using a 10% filter also left the data intact !
- ★ Average link cluster analysis also gives clustering by class



Matrix 2 - 10% Reduced



Summary

- ★ MALDI-TOF MS offers a rapid means of fingerprinting natural product extracts by molecular mass analysis
- ★ Data analysis techniques for clustering of mass spectra allow rapid discrimination of extracts
- ★ This information can be integrated into screening programmes to improve lead selection

