Smart Notes

How can the Orbitrap Exploris GC 240 mass spectrometer and Compound Discoverer software lead discovery in metabolomics studies?

For laboratories whose primary focus is metabolomic studies, the capability to comprehensively and accurately characterize and quantify the metabolome is critical. Gas chromatography high-resolution accurate mass (HRAM) mass spectrometry provides supportive information on volatile, semi-volatile, and derivatized polar metabolites across a range of biological, environmental, and clinical sample matrices. This capability in conjunction with liquid chromatography-mass spectrometry provides the ultimate coverage into the metabolome. A complete process to generate meaningful conclusions also includes careful experimental design and conduct prior to sample preparation. The Thermo Scientific™ Orbitrap Exploris™ GC 240 mass spectrometer delivers the very highest data quality for metabolomics to enable researchers to gain unprecedented depth of analysis and to have the flexibility to adapt to evolving research demands, such as biomarker discovery or metabolic pathway identification. With this analytical versatility, the system can provide the deepest insight into metabolomic studies.

ThermoFisher SCIENTIFIC

High quality raw data is important, but equally so, are informatics solutions to take that data and make clear discoveries between sample groups and to follow up with confident compound identifications resulting in actionable outcomes in an easy-to-use platform. The unrivalled combination of high-sensitivity full-scan HRAM data and intelligent identification using Thermo Scientific™ Compound Discoverer™ software addresses these challenges to advance scientific understanding. In this Smart Note, we take a closer look at metabolomics with the Orbitrap Exploris GC 240 MS[1] and Compound Discoverer software.

Metabolomics aims to characterize and quantify the complete small molecule complement, or metabolome, of a biological system. The metabolome consists of a diverse mixture of small molecules, including amino acids, sugars and phosphosugars, and biogenic amines and lipids. Untargeted metabolomics is exceptionally challenging due to the requirement to both identify and quantify hundreds of different compounds with limited a *priori* knowledge of the metabolites. Therefore, it is advantageous to use a detection system that is not only capable of sensitive detection of specific molecules in an untargeted way but can also provide accurate mass information and spectral information for confident confirmation and structural elucidation of unknowns. Gas chromatography-mass spectrometry (GC-MS) is routinely used for metabolomics applications due to its inherent advantages, especially its easy method development, chromatographic resolution, reproducibility, peak capacity, and extensive commercially available spectral libraries. GC provides excellent chromatographic separation capability for biomarker discovery using untargeted metabolomics. However, it has previously been hampered by the lack of high-end mass spectrometry support providing the dynamic range, accurate mass, and scan rate sufficient to analyze complex biological samples. The polar nature of most central metabolites means that derivatization must be performed to allow effective volatilization and ensure good chromatography. High sample throughput and advanced automation is required for metabolomic analysis, especially for clinical metabolomics. GC-MS using an Orbitrap-based mass spectrometry detector enables ultra-high mass resolution, sub-ppm mass accuracy, a large dynamic range, and a scan rate commensurate with the efficient quantitative analysis of highly complex metabolomic samples. The high resolution, mass accuracy, and scan speed is critical for consistent data deconvolution to permit the detection of species from overlapping TIC peaks. Accurate mass electron ionization (EI) fragment patterns are also suitable for matching against the widely available NIST and Wiley libraries for tentative compound identification, while providing accurate mass for more in-depth characterization.

Metabolomics with Thermo Scientific Orbitrap Exploris GC 240 MS and Compound Discoverer software

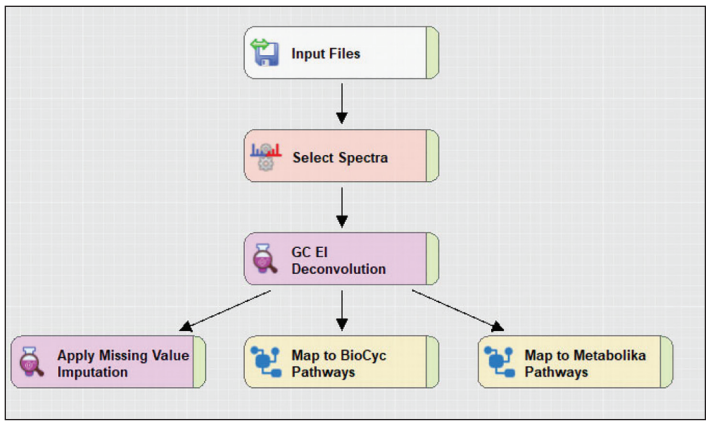
A well-planned untargeted metabolomics study begins with careful study design aimed at eliminating systematic and experimental bias. Second to this comes sample handling, metabolite extraction, and preparation prior to instrumental analysis. As described previously, a key aspect to a successful study is high quality accurate mass data that is processed using a powerful and fast informatics platform. Compound Discoverer software provides this capability to transform high quality data into meaningful results with ease. Starting with feature extraction, statistical analysis, compound identification, and biological pathway analysis, this combination allows researchers to obtain an unprecedented depth of analysis, providing a route to robust scientific discoveries.

Figure 1 details the instrumental and data processing workflow using the Orbitrap Exploris GC 240 MS. Data is first collected using EI in full scan at mass resolving power of 60,000 (FWHM *m/z* 200) and processed in Compound Discoverer software using the EI workflow node. This node performs feature extraction by deconvoluting the data to generate a list of features across all samples. Statistical analysis can also be carried out to identify differences between sample groups or individual samples. For example, these could be differences of study variables such as treatment type, age, origin, etc. These features, once filtered as being statistically significant (a potential biomarker), can then be matched against a nominal mass spectral library such as NIST or Wiley to make a putative compound identification. Further analysis could also be performed using a high-resolution accurate mass library for example Thermo Scientific™ Orbitrap™ GC-MS HRAM Metabolomics Library. Further confirmation of the library hit can be obtained by analysis using positive chemical ionization (PCI) on the Orbitrap Exploris GC 240 MS. This additional confirmation in the annotation of putative compounds and metabolites can be achieved by assessing the PCI spectra to predict the elemental composition of the molecular ion by looking at common adducts associated with a reagent gas. In PCI experiments using methane as the reagent gas, three adducts are typically observed: $[M+H]^+$, $[M+C_2H_5]^+$, $[M+C_3H_5]^+$. Alternative PCI reagent gases can be used including isobutane and ammonia.

Figure 1. Metabolomics workflow using the Orbitrap Exploris GC 240 MS and Compound Discoverer software, taking high quality data through to confident metabolite identification and confirmation

Figure 2A shows the details of the two nodes available within the Compound Discoverer software specific for processing Orbitrap Exploris GC EI and PCI data. The EI node includes an initial peak deconvolution to extract the peak features within a data set. These peak features can then be statistically compared to other sample groups prior to compound annotation. The first step in compound identification is to match the clean deconvoluted spectra to existing spectral libraries (Figure 2B). Within the node these can be specified as nominal mass libraries such as NIST or high-resolution accurate mass libraries. The PCI node also included peak feature extraction, but instead of matching to spectral libraries, the molecular ion is identified through mass adducts, and elemental compositions are proposed for the molecular ion using a pre-defined set of elements. With excellent sub-1 ppm mass accuracy expected from Orbitrap MS, the list of possible elemental composition is small, enabling the correct assignment to be made and within a short time frame. Once an elemental composition is proposed, this can then be automatically searched against ChemSpider online database of compounds to make a putative annotation. Absolute confirmation of compound identify will require comparison of retention time and spectra against a certified standard.

GC EI workflow



GC PCI workflow

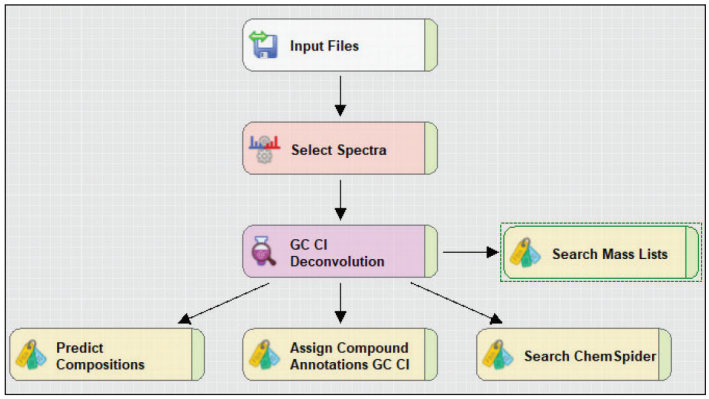


Figure 2A. GC workflows in Compound Discoverer software, covering both EI and PCI HRAM data

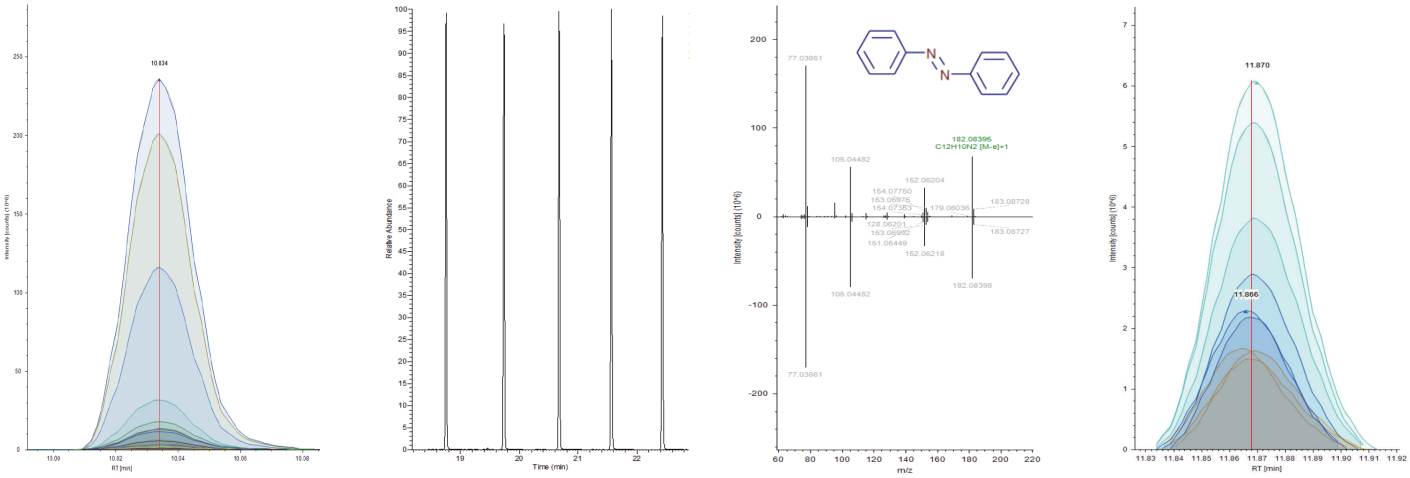


Figure 2B. Electron ionization node workflow within Compound Discoverer software

In the results browser of Compound Discoverer software, all the peak detection, annotation, and statistical tools are easily available and clearly displayed to the analyst for review. Figure 3 shows the results browser with EI spectral match results to a library entry. Library search index score, ppm mass accuracy of molecular ion, retention index, and high-resolution filtering score are displayed alongside the mirror plot of the library top hit. The extracted ion chromatogram of the compound is also displayed to enable the user to quickly see how this peak intensity varies across all samples within the analytical batch. The sensitivity and selectivity of the Orbitrap Exploris GC 240 MS enables the software to work efficiently, and high quality results with good confidence are displayed to the analyst in this report. For example, in this case the compound is putatively annotated as camphene with 0.7 ppm mass accuracy, a search index of 837 and HRF score 95.3, and retention index delta of 1.

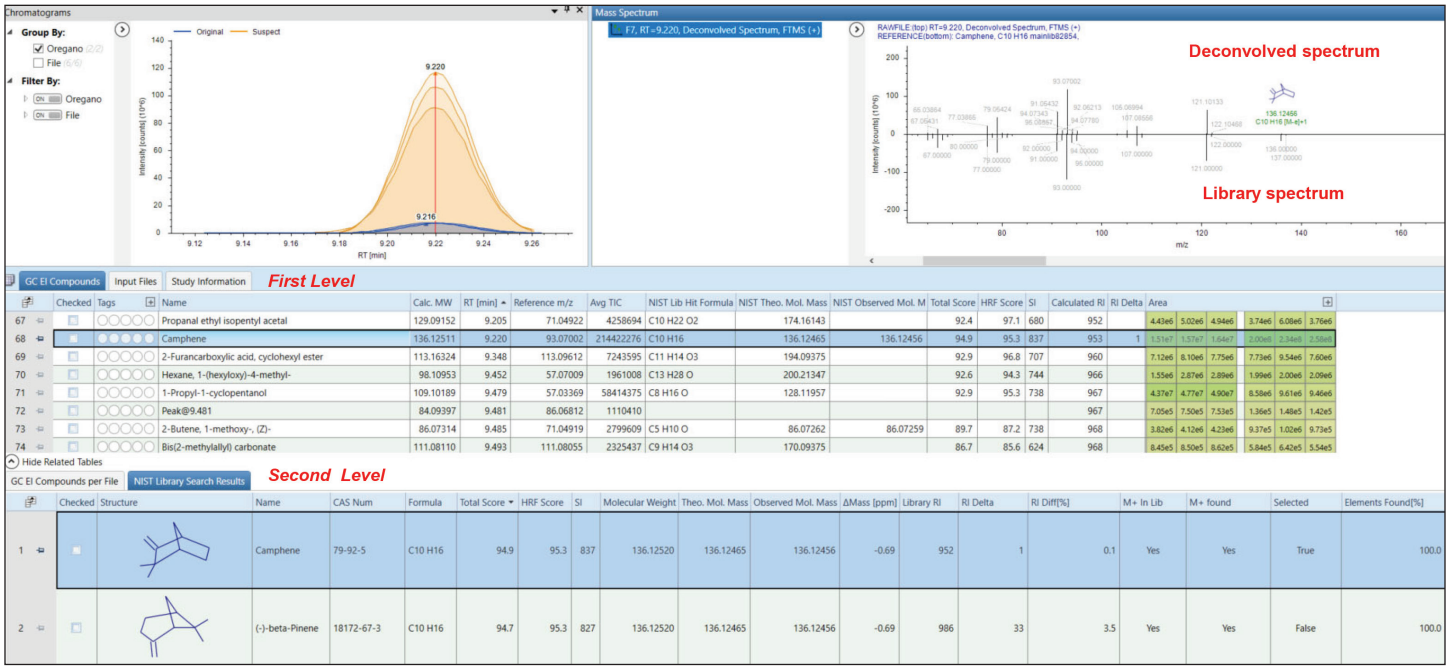


Figure 3. Compound Discoverer library search result browser view. Display shows extracted ion chromatogram across all samples, library match mirror view, list of peaks, and identification results.

Orbitrap GC-MS HRAM Metabolomics Library

With the need to decipher biological complexities, metabolomics analysis demands sophisticated analytical technologies and advanced software solutions. The Orbitrap GC-MS HRAM Metabolomics Library is the first commercially available high-resolution accurate mass metabolomics library for electron ionization GC-MS. It contains more than 900 retention indexed unique entries from more than 800 metabolites, providing a broad coverage of primary and secondary metabolites (including volatiles) in plants, animals, and microbes (Figure 4). When used in combination with powerful Orbitrap GC-MS technology and unique Thermo Scientific software data processing tools, the difficult challenge of metabolite identification in untargeted experiments is easier than ever before. Library entries, many with methoxyamine/MSTFA derivatization were acquired at 60,000 mass resolving power and include Kovats retention index information for additional confirmation.

Positive chemical ionization to annotate the molecular ion

A complimentary approach is to identify unknown compounds or confirm spectral match results using PCI spectral data within Compound Discoverer software. Using the PCI node, the data is deconvoluted to determine peak features across all samples, and these can also be compared to each sample group as shown in Figure 5. The focus in PCI data is to identify molecular ions using expected mass adducts to the molecular ion.

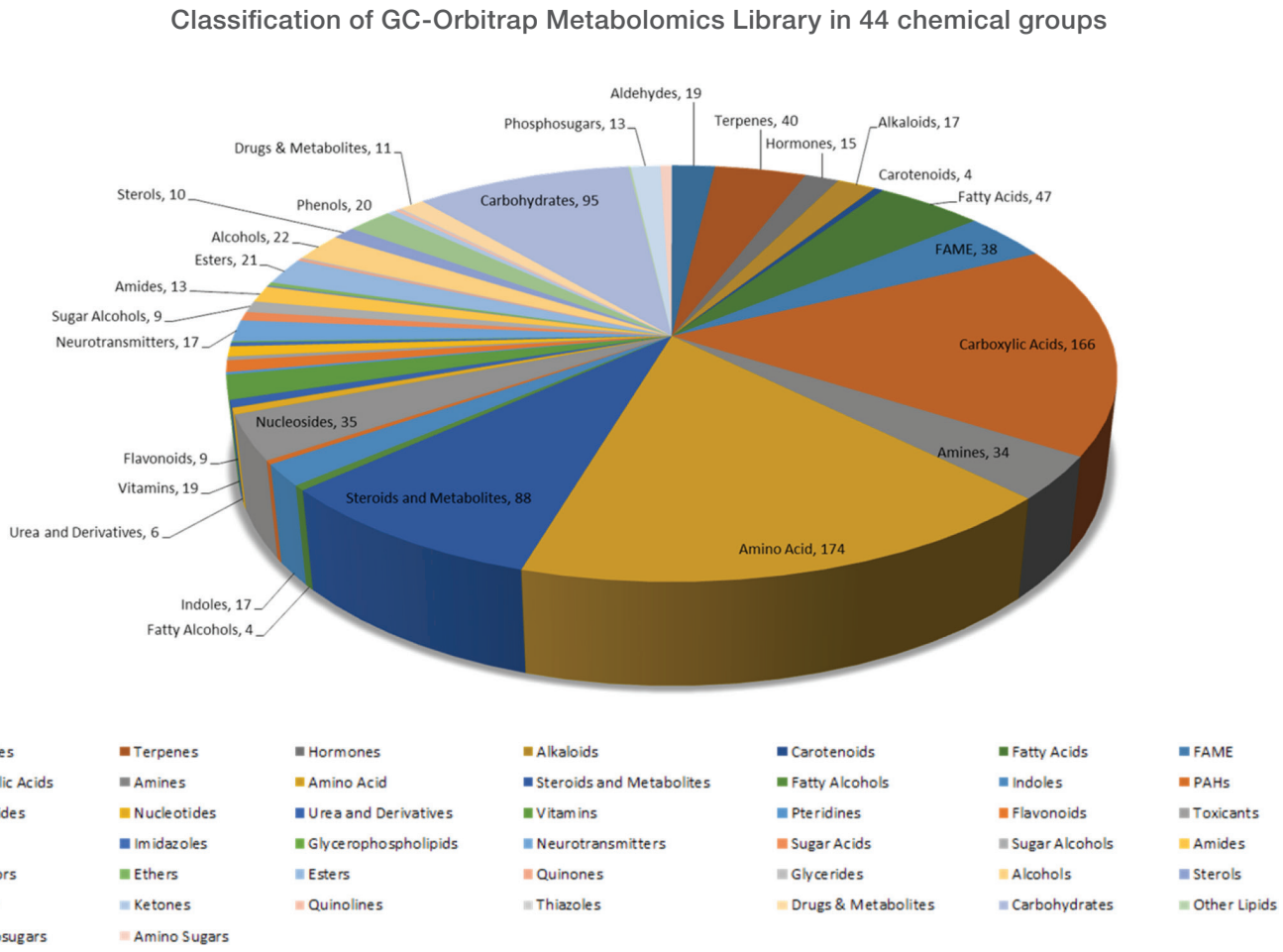
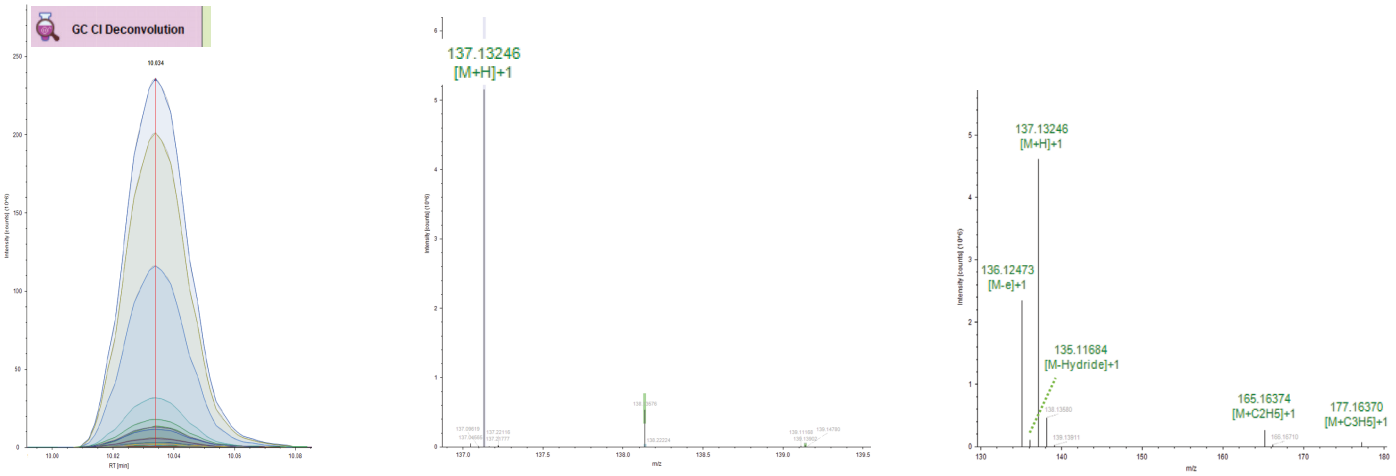


Figure 4. Orbitrap GC-MS HRAM metabolomics library compound contents details



Peak deconvolution | Molecular ion determination | Adduct assignment

Figure 5. Chemical ionization node functionality with the aim of identifying molecular ion adducts to support unknown identification or compound confirmation from spectral match result. Peak deconvolution is followed by automatic assignment of molecular ion adducts (center and right).

Statistical comparison information

Multivariate statistical analysis is easily accessible within Compound Discoverer software alongside interactive annotation and filtering tools. Figure 6 provides examples of some classical approaches to detect differences between sample groups including PCA, hierarchical cluster, box plots, and volcano plots. PCA is a well-known statistical approach that highlights variation between sample groups and allows visualization of strong patterns in complex datasets. The first step will inform if there are any differences between the sample groups. Ideally, a pooled sample will have been included in the data acquisition, and this should fit around the center point of the PCA to validate the experiment. With confirmation of significant statistical differences, the next action is to determine which peak or peaks contribute to that difference, identifying the markers unique to a sample group. The volcano plot is a powerful tool to identify the features that are unique to a group, and from here the compounds can be identified using the approaches described above.

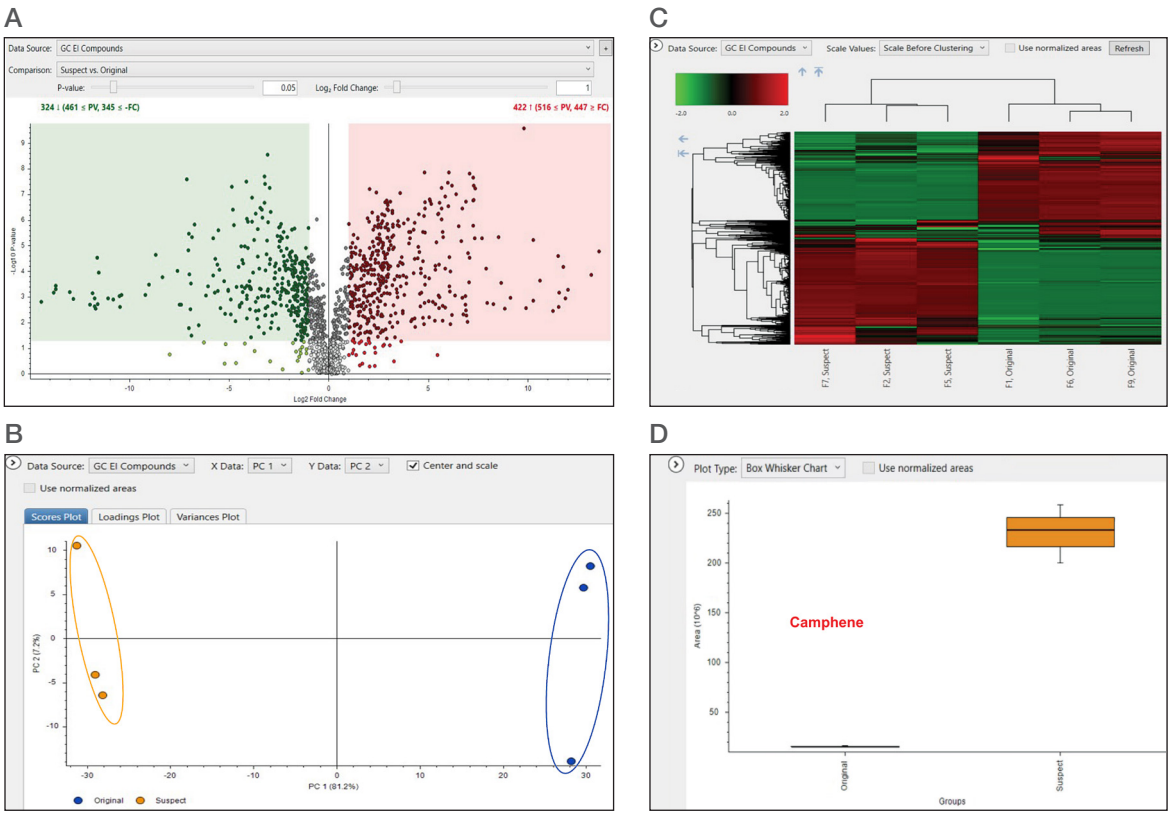


Figure 6. Statistical analysis tools in Compound Discoverer software with EI data. (A) Volcano plot, (B) PCA, (C) Hierarchical cluster analysis, (D) Boxplot

Compound Discoverer software will automatically assign $[M+H]^+$, $[M-H]^+$, $[M+C_2H_5]^+$, $[M+C_3H_5]^+$ ions in the spectrum during the deconvolution. If instructed in the workflow, the software will predict the elemental composition of the molecular ion based on these adducts and the elements included within the search parameters. All the results and matched criteria are displayed within the results browser (Figure 7). With sub-ppm mass accuracy, the number of suggested elemental compositions is low and the confidence in assignment high. The end result is that researchers gain unprecedented depth of analysis and speed to result.

For true unknowns with PCI MS^2 scan, the elemental composition will be screened against the ChemSpider database. Compound Discoverer software calculates a FISh score to generate in silica MS^2 fragmentation spectra for comparison between the acquired ions and the theoretical fragmentation (Figure 8). The fragments are matched with theoretical pattern and annotated in green. The higher the FISh scores, the higher the confidence in the assignment.

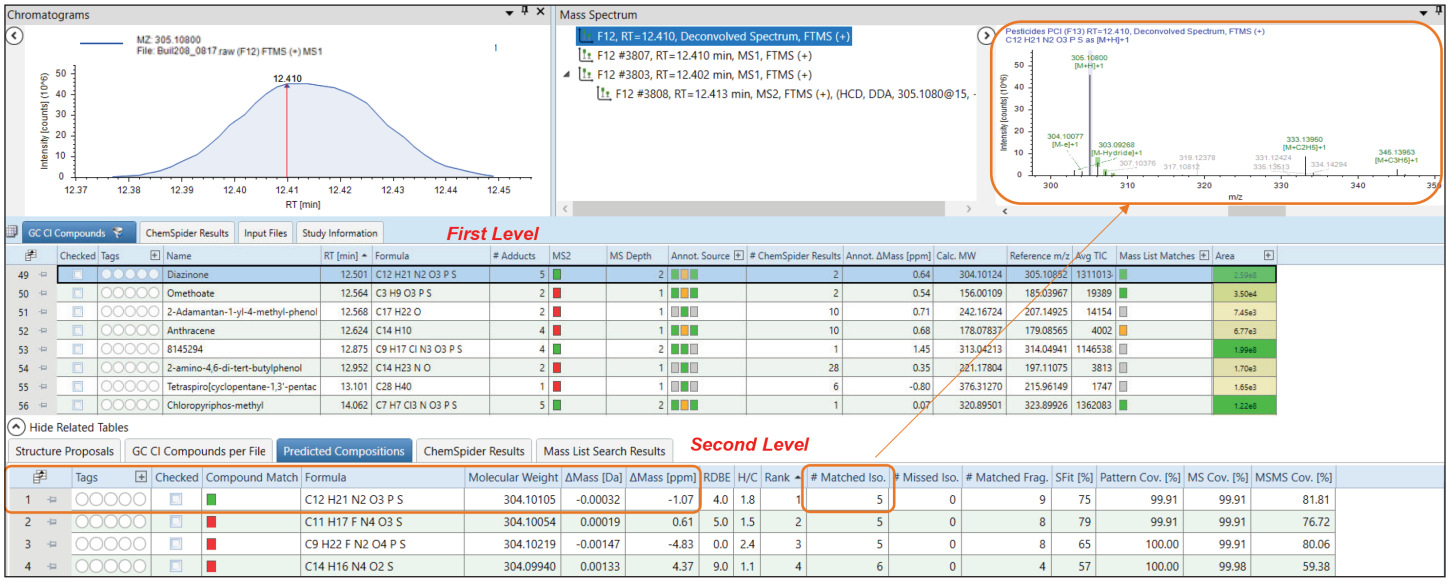


Figure 7. GC CI results. Predicted elemental compositions based on accurate mass and isotope pattern. The molecular ion is identified through adduct assignment and elemental compositions proposed and supported with theoretical isotope pattern matching.

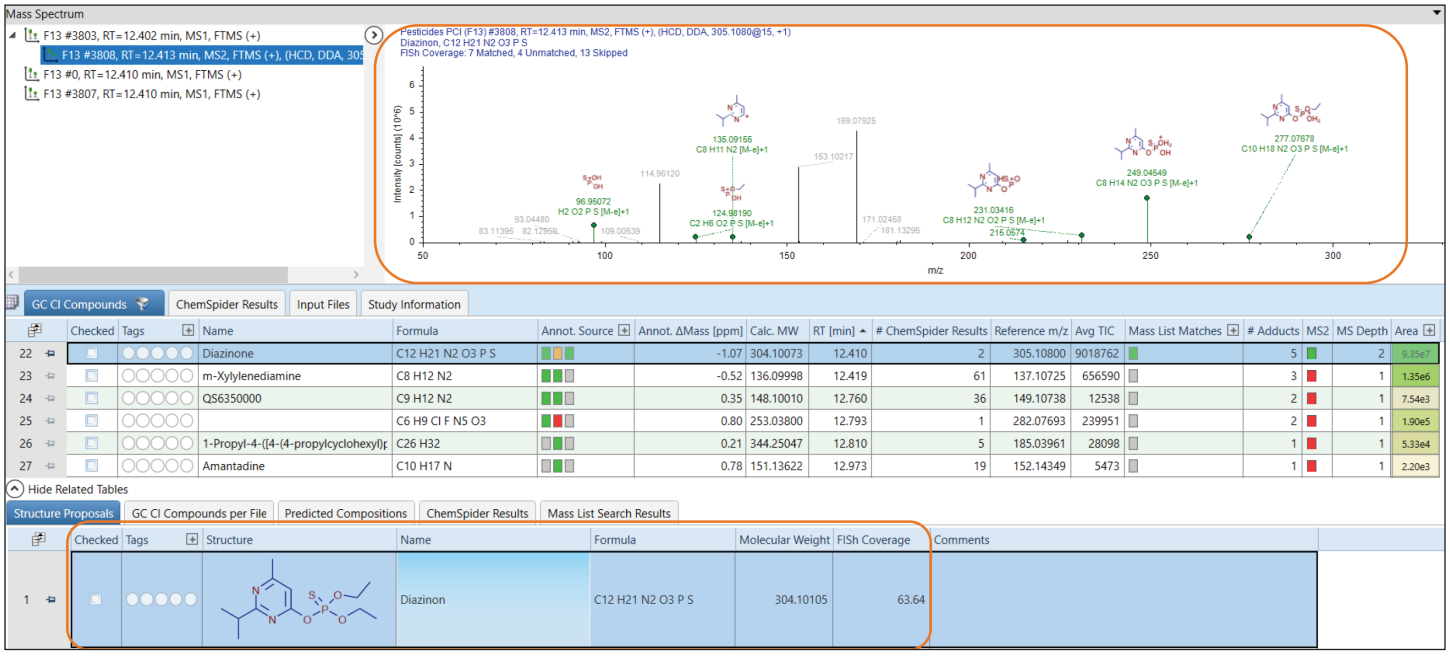


Figure 8. GC CI results. FISh Scoring for MS^2 scan. Identification through online database matching is performed and fragments assigned and annotated to mass spectrum.

thermo scientific

Summary

The data quality from the Orbitrap Exploris GC 240 MS combined with the functionality of Compound Discoverer software demonstrate a powerful tool for metabolomic studies. The solution provides an integrated metabolomics approach for targeted analysis, profiling complex samples, and identifying unknown compounds.

- The streamlined GC-EI and GC-CI data processing workflow nodes in Compound Discoverer software allow for extraction, deconvolution, identification of unknown compounds, and multivariate statistical and pathway analysis.
- Principal components and differential analysis tools allow for the detection of differences in the composition of sample groups. After peaks are isolated, they can be quickly screened against existing nominal mass or high-resolution accurate mass spectral libraries.
- The high resolving power, consistent sub-1 ppm mass accuracy, and wide dynamic range allow for fast and confident characterization of compounds, regardless of their concentration or matrix complexity.
- Rapid change-over from EI (for spectral library search) to softer ionization such as PCI (for molecular ion confirmation using adduct information) and the ability to perform accurate mass MS/MS experiments (for structural elucidation), enable confident compound identification with unprecedented ease.

Reference

1. Thermo Fisher Scientific Application Note 10724: Breakthrough performance of the Orbitrap Exploris GC for analytical testing and scientific research applications, 2020. https://assets.thermofisher.com/TFS-Assets/CMD/Application-Notes/an-10724-gc-ms-breakthrough-performance-orbitrap-exploris-an10724-en.pdf

Find out more at
thermofisher.com/OrbitrapExplorisGC240

ThermoFisher SCIENTIFIC