

## Application of Multivariate Analysis and LC-MS for the Detection of Counterfeit Cosmetics

---

Baiba Cabovska

日本ウォーターズ株式会社

---

### Abstract

In this application note, we describe a rapid, simple method using high-resolution mass spectrometry and multivariate analysis to compare authentic cosmetics samples with suspected counterfeit samples. This LC-MS and informatics workflow can be adopted for cosmetics, food and beverage, and pharmaceutical sample analysis.

### Benefits

- A rapid, simple method using high-resolution mass spectrometry and multivariate analysis to compare authentic cosmetics samples with suspected counterfeit samples.
- Can be adopted for comparative analysis of cosmetic products samples, as well as for other types of analysis where an evaluation of differences is needed (e.g. failed batch of raw materials or packaging).

---

### Introduction

Every year the cosmetics industry suffers multi-billion dollar losses due to counterfeit cosmetic products.<sup>1</sup> Due to much stricter regulatory controls in Europe and North America, 66% of counterfeit goods come from

Asia.<sup>2</sup> The risk to the consumer is high, not because they are paying for a counterfeit product, but because the ingredients used in the production of counterfeit cosmetics could be harmful to their health, or even banned for human use.

Testing for counterfeit products is occasionally done by the cosmetics companies – especially those companies whose high-end products are usually the target of counterfeiting, since only they know the correct formulation of their products. However, it would be beneficial and less time consuming if counterfeit testing could be done at the point of entry in the country, for instance, during customs inspection. Even if the correct formulation is not known, it is possible to compare suspected fake samples with authentic samples using multivariate statistical analysis and assess the differences if needed.

Multivariate analysis (MVA) is widely used in the areas where multiple samples or batches need to be compared. One of the most commonly used techniques is principal component analysis (PCA) which allows the reduction of a large set of multivariate data into uncorrelated variables called principal components.

---

## Experimental

### Sample preparation

High-end cosmetic samples were purchased from the U.S. manufacturer. A cream, a lotion, and a serum were chosen for this study. Identical looking items were resourced from an online retailer in Asia. All samples were prepared by dilution in tetrahydrofuran (THF) at a concentration of 5 mg/mL.

### UPLC conditions

UPLC system:	ACQUITY UPLC I-Class
Separation mode:	Gradient
Column:	CORTECS UPLC C <sub>18</sub> , 90Å, 1.6 µm, 2.1 mm x 100 mm
Column temp.:	40 °C
Injection volume:	5 µL
Flow rate:	0.4 mL/min

Mobile phase A:	0.1 % formic acid in water
Mobile phase B:	0.1 % formic acid in methanol
Gradient:	20% B held for 30 s, increased to 99% over 2.5 min, held at 99% for 6 min, then re-equilibrated back to 20%

## MS conditions

MS system:	Xevo G2-XS QTof
Ionization mode:	ESI + and -
Capillary voltage:	3.0 kV for pos, 2.0 for neg
Desolvation temp.:	450 °C
Source temp.:	150 °C
Cone voltage:	30 V
Collision ramp:	10 to 40 eV
MS scan range:	50 to 1200 <i>m/z</i>

## Data acquisition and processing

Waters UNIFI Scientific Information System was used for data acquisition and processing.

---

## Results and Discussion

Samples were analyzed as five replicates in positive and negative ESI mode to obtain a representative data set in both ionization modes. Due to the lack of information about the counterfeit sample formulations, and for a more comprehensive analysis, data was acquired using different ionization modes as some compounds will ionize exclusively with positive ionization, and others only by negative ESI.

All sample data was processed using the multivariate analysis tools available in UNIFI Scientific Information System. UNIFI can generate marker matrices based upon user-defined criteria that can be automatically transferred to EZInfo software for MVA. The initial summary is presented as a PCA scores plot. In this initial plot no information about the individual sample groups is passed to the MVA software, and this model is said to be unsupervised.

If additional discrimination among the investigated sample groups is required, a supervised analysis model, such as the Projection to Latent Structures Discriminant Analysis (PLS-DA) model (Figure 1) can be employed. PLS-DA models the quantitative relationships between the variables X (predictors) and Y (responses) for all the sample groups and can be used to elucidate group differences. However, in these types of plots, each sample is presented by a single point, which does not allow individual markers contributing to the differences between the groups to be observed.

In Figure 1a, the data obtained by ESI- is presented. It can clearly be seen that there are differences between each of the samples. In this plot a general trend can be observed that the authentic product samples fall in the lower quadrant, biased toward the left side, while the counterfeit samples appear at the top and toward the right.

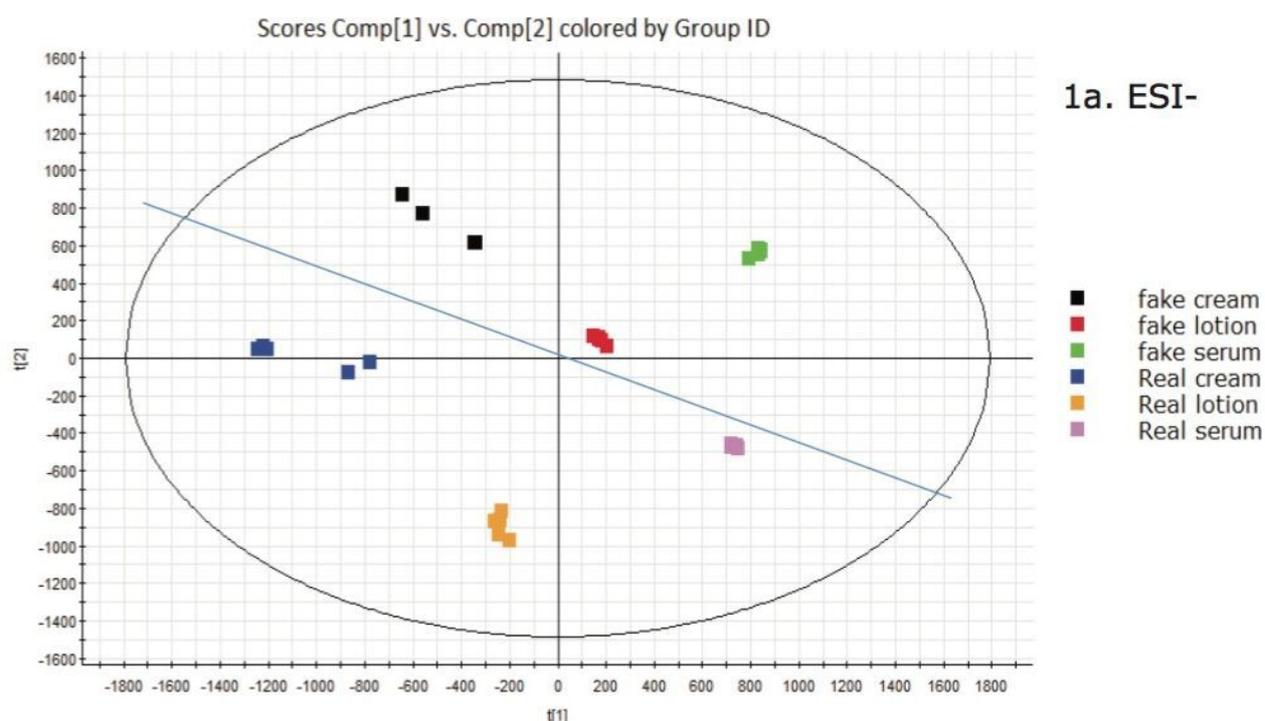


Figure 1a. PLS-DA plot for all the samples in ESI- mode.

For clarity, a blue line has been added to the plot showing the counterfeit data above the blue line, and all the authentic product samples below it. This plot indicates that there is some element of the data that is

common in the ESI- and contributes to the grouping of the two sets of samples. Although significant differences are also seen in positive ion mode (Figure 1b), no general trend for counterfeit versus authentic samples was observed.

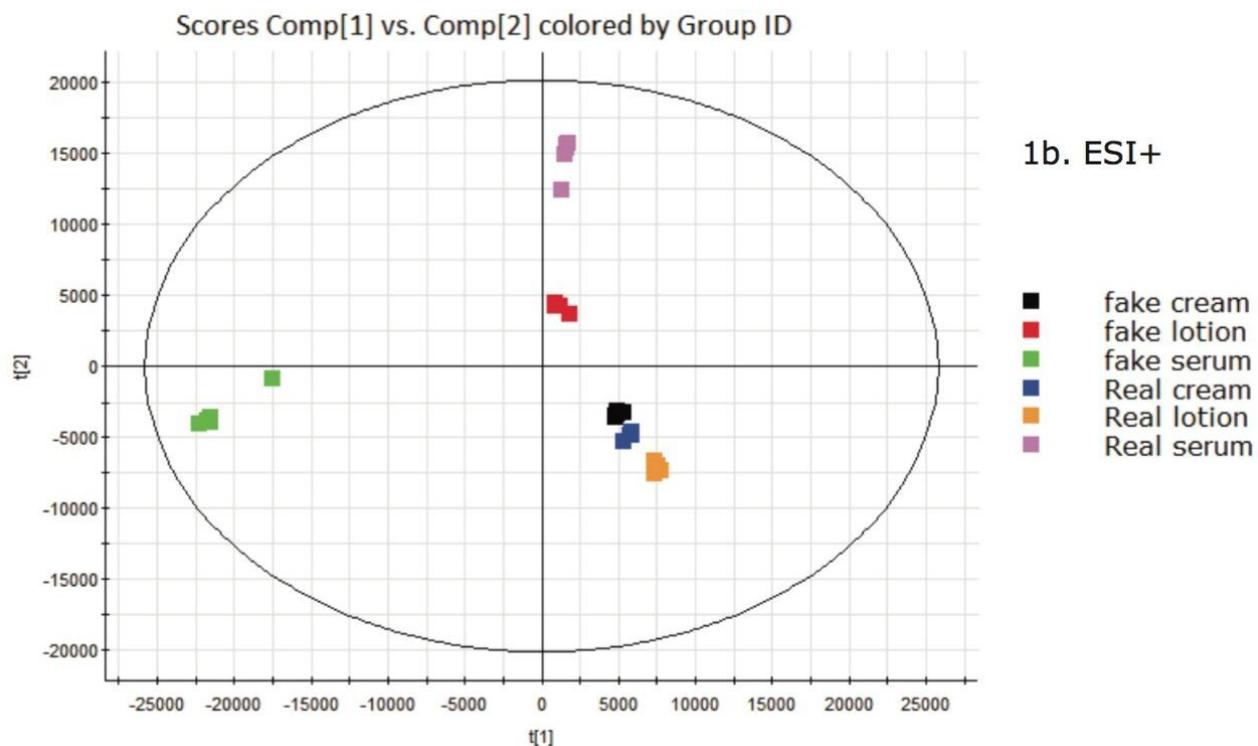


Figure 1b. PLS-DA plot for all the samples in ESI+ mode.

Each of the sample groups (creams, lotions and serums) were further investigating by using Orthogonal Projection to Latent Structures – Discriminant Analysis (OPLS-DA) scores plots, shown in Figure 2.

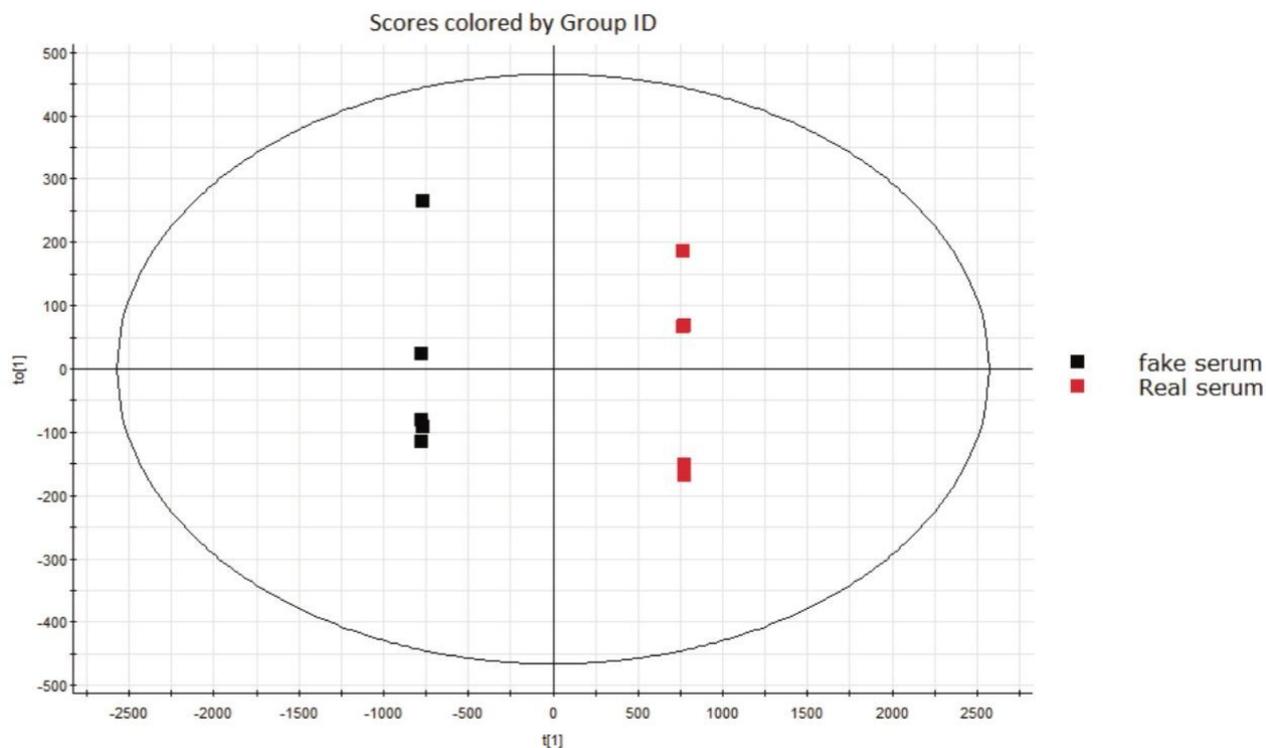


Figure 2. OPLS-DA plot for serum samples in ESI- mode.

OPLS allows analysts to mine the data for additional information beyond that of simple differences between groups. This additional level of detail is needed to identify specific features of the data that contribute to what makes the samples different from one another, such as to discover whether the difference in the counterfeit samples is due to chemicals that are harmful to human health. The tool used to dig deeper into the data is called an S-plot. The S-plot shows the Accurate Mass/Retention Time (AMRT) dissimilarities between these two groups, shown in Figure 3.

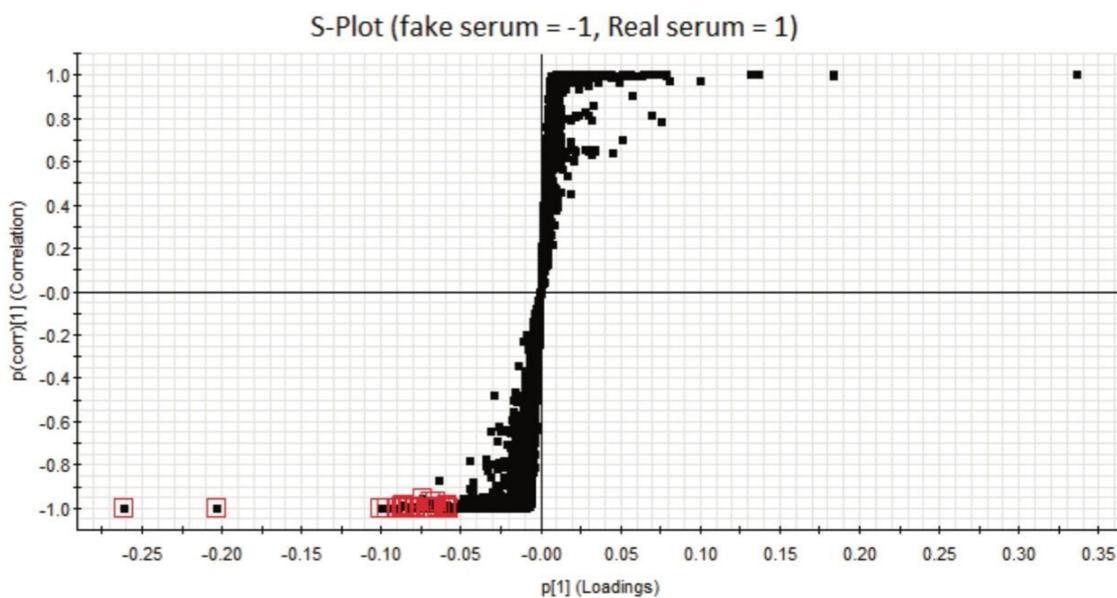


Figure 3. S-plot for counterfeit and authentic serum samples in ESI- mode. Markers selected in red have the greatest contribution to the variance between the fake serum and the authentic one.

The AMRT pairs are plotted by covariance – the magnitude of change (x-axis), and correlation – the consistency of the change (y-axis) values. The upper right quadrant of the S-plot shows AMRTs which are elevated in the authentic sample, while the lower left quadrant shows components elevated in the counterfeit sample. In this case, an AMRT may represent a component of the formulation that is different between the two samples. The farther along the x-axis the marker is located, the greater its contribution to the variance between the groups, while markers farther along the y axis represent a higher reliability of the analytical result. The differences between the groups can come from analytes which are not present in one of the groups, or from analytes with the greatest change in intensity (concentration) between the groups.

After selecting the markers which contribute to the differences between the groups, each marker set can be labeled accordingly. For example, markers specific to the differences between fake cream and the authentic cream. The labels can be appended if the marker is found in more than one selection. Such group comparison was done for all three types of samples. After labeling each group, it was observed that two markers were present in all three types of counterfeit samples –  $m/z$  151.0409 at 2.59 mins, and  $m/z$  179.0725 at 3.64 mins (Figure 4). These two AMRTs very likely contributed towards the distinct separation observed in the PLS-DA (ESI-) plot between the authentic samples and the counterfeits. In the trend plot, it was also observed that these markers were not detectable in the authentic samples.

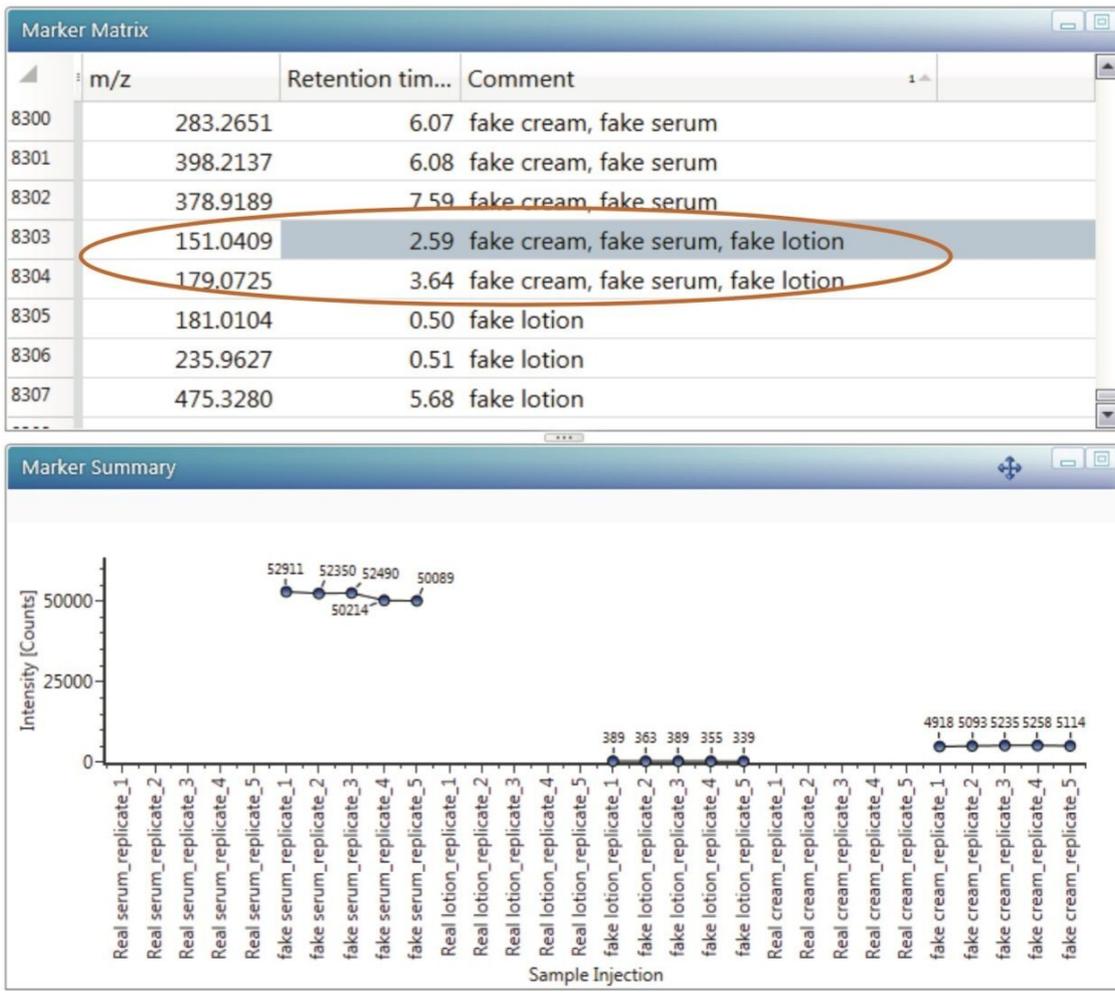


Figure 4. Marker table and the trend plot for ESI- data.

To investigate the markers further, the discovery tools in the UNIFI Scientific Information System were employed. Both markers were submitted for automated elemental composition calculation, structural database search, and fragment matching of the high collision energy data. The results are shown in Figure 5.

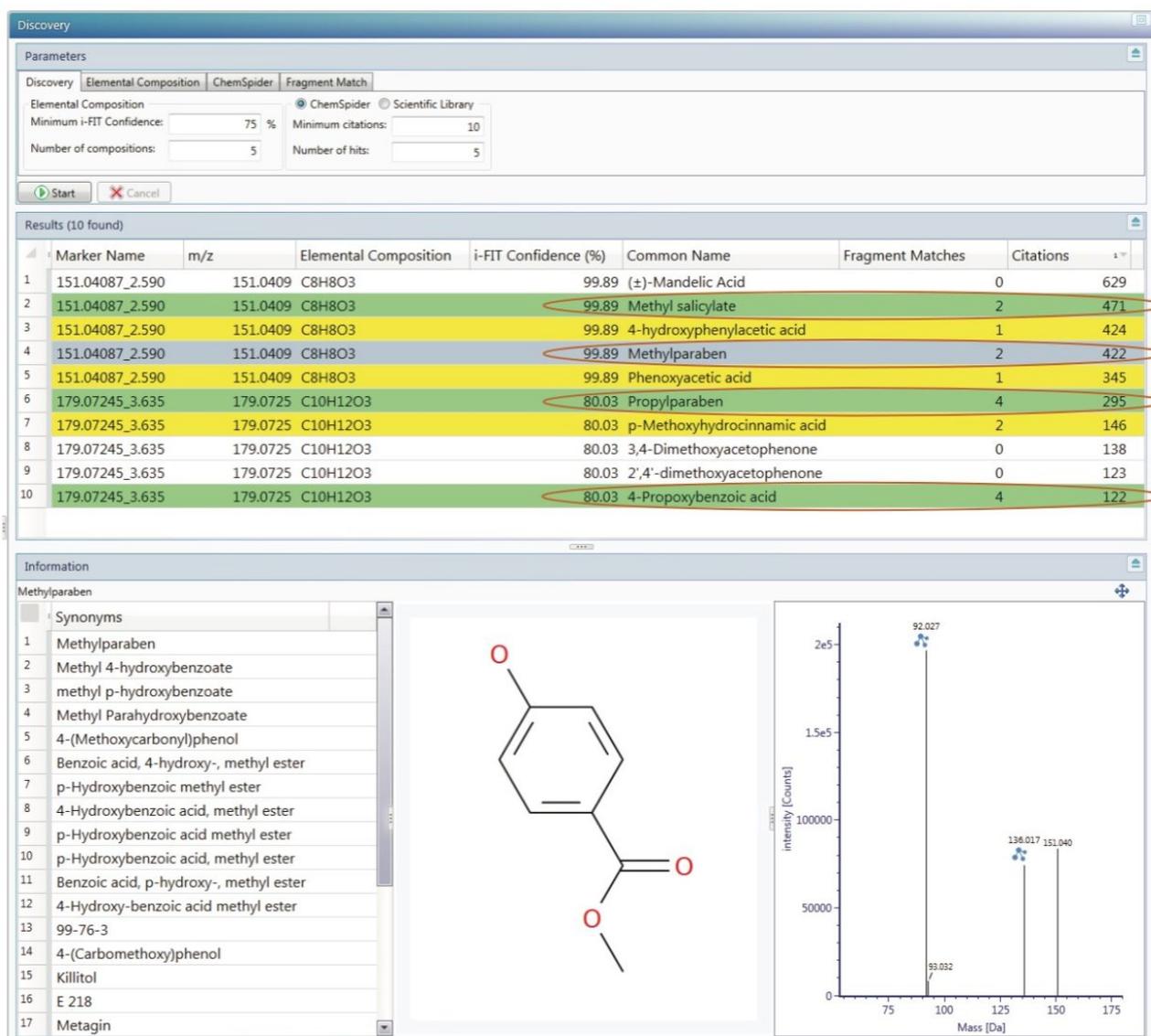


Figure 5. Discovery tool results for m/z 151.0409 and 179.0724. Summary of elemental compositions, citations for structures retrieved from ChemSpider database and a number of possible fragment matches in high collision energy data for each structure.

For the first marker a molecular formula of  $C_8H_8O_3$  was proposed. From the corresponding structures in the ChemSpider database, methylparaben and methyl salicylate have most high collision energy fragments matched (2). The second marker has a molecular formula of  $C_{10}H_{12}O_3$  and propylparaben and 4-propoxybenzoic acid have the most fragments matched (4). Methyl salicylate is used as a fragrance in foods and beverages. 4-propoxy benzoic acid can be used in chemical synthesis of liquid crystals. Instead, parabens are preservatives most commonly used in personal care products like body lotions and creams.<sup>3</sup> However, due to public awareness and concerns about parabens being endocrine disruptors, high-end

cosmetics companies have stopped using them in their products.<sup>4,5</sup>

The producers of the counterfeit products have no such concern, and in this case, appear to have formulated the product using the lowest cost chemicals available to obtain a product which superficially appears the same as an authentic sample.

Based on information available from the data and discovery tools, together with the information available on the chemicals proposed, there is high confidence in the assignment of parabens to these two markers at *m/z* 151.0409 and 179.0724.

---

## Conclusion

Every year the cosmetics industry suffers from multi-billion dollar losses due to counterfeit cosmetics products. This lost revenue may have a negative impact on market share, and can result in a further erosion of sales. If the counterfeit products cause health problems in consumers, this can damage the reputation and brand image for the manufacturers of the authentic cosmetics. Early and rapid detection of counterfeit products is one way to address counterfeiting in both domestic and export markets. Highlighted in this work is a multivariate analysis technique for sample comparison using statistical analysis tools for easy comparison between complex samples. The described LC-MS and informatics workflow as implemented with the UNIFI Scientific Information System using high-resolution mass spectrometry can be adopted for cosmetics, food and beverage, and pharmaceutical sample analysis.

---

## References

1. <https://oami.europa.eu/ohimportal/en/web/observatory/news/-/action/view/1934074>
2. [www.ccapcongress.net/archives/Brussels/Files/fsheet5.doc](http://www.ccapcongress.net/archives/Brussels/Files/fsheet5.doc)
3. <http://www.fda.gov/Cosmetics/ProductsIngredients/Ingredients/ucm128042.htm>
4. <http://www.businessinsurance.org/the-truth-about-cancer-causing-cosmetics/>
5. <http://fitbeaut.blogspot.com/2012/10/paraben-free-cosmetics.html>

---

## Featured Products

ACQUITY UPLC I-Class System <<https://www.waters.com/134613317>>

UNIFI Scientific Information System <<https://www.waters.com/134801648>>

Xevo G2-XS QToF Mass Spectrometer <<https://www.waters.com/134798222>>

720005402, May 2015

©2019 Waters Corporation. All Rights Reserved.