Egg product freshness evaluation: A metabolomic approach

Daniele Cavanna1,2 | Dante Catellani1 | Chiara Dall'Asta2 | Michele Suman1

1 Advanced Laboratory Research, Barilla G e R. Fratelli S.p.A., Parma, Italy
2 Department of Food and Drug, University of Parma, Parma, Italy

Correspondence
Dr Michele Suman, PhD, Barilla G. R. F.li SpA, Advanced Laboratory Research, via Mantova 166, Parma 43122, Italy.
Email: michele.suman@barilla.com

Abstract
Egg products' freshness is a crucial issue for the production of safe and high-quality commodities. Up to now, this parameter is assessed with the quantification of few compounds, but the possibility to evaluate more molecules simultaneously could help to provide robust results.

In this study, 31 compounds responsible of freshness and not freshness of egg products were selected with a metabolomic approach. After an ultrahigh-pressure liquid chromatography–high-resolution mass spectrometry (UHPLC-HRMS) analysis, different chemometric models were created to select gradually the most significant features that were finally extracted and identified through HRMS data.

Sample lots were collected directly from their arrival at the production plant sites, extracted immediately after, then left at room temperature, and extracted again after 24 and 48 hours (first day and second day, respectively). A total amount of 79 samples was used for the model creation.

Furthermore, the same compounds were detected in seven new egg products sample lots not used for the model creation and treated with the same experimental design (total amount of samples, 21).

The results obtained clearly demonstrate that these 31 molecules can be considered real freshness or not freshness chemical markers.

Furthermore, this UHPLC-HRMS metabolomic approach allows for the detection of a larger set of metabolites clearly related to possible microbial growth over time, which is a relevant point for also ensuring food safety.

KEYWORDS
egg product, freshness, LC-HRMS, metabolomics, nontargeted mass spectrometry

1 | INTRODUCTION

Starting from the beginning of the new millennium, consumers' attention on the authenticity and quality of food commodities has strongly increased, leading to an increase in the demand of fit-for-purpose methods for detecting food fraud from both industries and research institutes.

Eggs, mostly in the egg products form, are largely used for the formulation of food products. Freshness is therefore a crucial parameter for ensuring the production of safe and high-quality commodities. However, the assessment of egg freshness is still challenging, due to the lack of robust chemical markers.

In the past, according to the European legislation, two parameters typical of microbiological fermentation had to be monitored: lactic acid...
by the current laws consider only lactic acid as reliable marker.  

These two compounds could act sometimes as “tardive” markers of eggs ageing, and their increase is often not linear during the time; moreover, the use of only one/two molecule(s) for the evaluation of such critical parameter can be not sufficient in some circumstances; at the same time, more “not freshness” compounds should be monitored.

Up to now, major attention was paid to the development of rapid techniques able to assess egg products freshness, eg, using electronic noses or spectroscopic analyses. These techniques are particularly suitable for industrial screening purposes, due to the reduced time and costs.

However, the identification of robust “markers of freshness” (ie, compounds that decrease their intensity during the egg products storage) could strongly support the evaluation of egg products ageing, with a reduction of the risk for the final consumer.

In this context, the use of “nontargeted” methods based on a metabolomic approach—which are considered emerging methodologies for detecting food frauds—will offer the opportunity to identify and validate proper markers.

In consideration of the ageing process, chances in the volatile profile of egg products can be expected. Therefore, among possible analytical techniques, gas chromatography–mass spectrometry could be used for the identification of these changes, in agreement with previous applications in the field of food frauds, ie, tomato cultivars discrimination, honey authenticity, or geographical origin of saffron.

However, the nonvolatile fraction of food may be of great interest for the discrimination of ageing.

Scientific papers suggest that other rapid mass spectrometry approaches, as, for example, direct analysis real-time–mass spectrometry could lead to the characterization of raw materials, avoiding the chromatographic separation.

In any case, liquid chromatography–high-resolution mass spectrometry (LC-HRMS) is the technique of choice for the execution of a metabolomic study: Thanks to its versatility, a huge amount of information related to the nonvolatile profile can be extracted from the chromatographic fingerprints.

Literature presents applications of this approach for the detection of frauds related to a wide range of raw materials, from fruit juices to wheat and from extra virgin olive oil to spirit drinks.

Independently from the analytical technique used, the creation of robust chemometric models is a crucial step for the selection of new marker compounds responsible of the target fraud.

Based on what described above, this work presents a metabolomic study on egg products samples able to select and identify a group of new compounds that can be used as “freshness” or “not freshness” markers.

Liquid chromatography–high-resolution mass spectrometry was exploited for fingerprints recording; subsequently, robust data elaboration was executed, creating different chemometric models able to select some marker compounds that were then studied and identified.

The final step of this workflow was the validation of these molecules, with the aim to assess their value as chemical markers, independently from the chemometric model used to select them.

## 2 MATERIALS AND METHODS

### 2.1 Chemicals and sample description

Methanol, acetonitrile, formic acid (FA), and ammonium formate (AF) were purchased from VWR International, Ltd (Poole, UK).

Analytical standards of chlorophenolic, lactic acid, succinic acid, phenylactic acid, tyramine, uracil, asparagus, glutenos, guanosine, guanosine monophosphate, serine, uridine, and uridine monophosphate were purchased from Sigma-Aldrich (St Louis, Missouri).

Water was purified using a Milli-Q system (Millipore, Bedford, Massachusetts).

Egg products samples were directly collected from batches that daily arrive at production plant sites; to increase the variability of the model, the sampling period lasted 6 weeks, and samples coming from different batches, different suppliers, and with different amounts of carotenoids (in the range of tenths of milligrams per kilogram, authorized for improving appearance and correspondent consumer preference in finished products) were selected.

### 2.2 Experimental design

The first step of the adopted workflow, as suggested by US Pharmacopoeia, required the creation of chemometric models able to discriminate between fresh and not fresh samples; subsequently, the features responsible of this clusterization were selected, and a tentative of compounds identification was performed.

Finally, marker compounds identified were validated, that means that the target molecules were searched in new egg products batches subject to the same ageing applied during the model creation.

For the creation of the model, 29 egg products batches were collected and extracted immediately after the reception at the production plant site; subsequently, these raw materials were left at room temperature and extracted again after 1 day and after 2 days, according to the sampling plan described in Table 1. Summarizing, a global amount of 79 samples was used for the creation of the chemometric model.

For the validation of the marker compounds, seven new egg products batches were collected, extracted, and analyzed following the same experimental design described for the model creation: The total amount of samples was 21.

### 2.3 Sample preparation

A total of 300 μL of each egg product at each time point was spiked into a 1.5-mL tube, together with 30 μL of a 20-μg/mL chlorophenicol solution (used as internal standard) and 900 μL of a mixture of acetonitrile to water in the ratio 80:20 (stored at 2°C–8°C).

This mixture was vortexed for 1 minute, then stored for 5 minutes at −20°C, and subsequently centrifuged for 10 minutes at 14 000 rpm with a Rotina 380R (Hettic Lab Technology, Tuttinglen, Germany) maintained at 4°C.

The extract was filtered into a high-performance liquid chromatography vial with a 0.22-μm PolyTetraFluoroEthylene (PTFE) syringe filter (Phenomenex, Torrance, California) and stored at −20°C until analysis with LC-HRMS instrument.
Gradient conditions are the following: After 1 minute with 95% of mobile phase A (0.1% FA and 5mM AF in water) and 5% of mobile phase B (0.1% FA and 5mM AF in methanol), the percentage of solvent B increased to 95% in 24 minutes and then was maintained at this percentage for 5 minutes before column re-equilibration (10 min).

Autosampler was maintained at 5°C, and the injection volume was 5 μL.

Mass spectrometry detection was performed with a benchtop Q Exactive Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, Massachusetts) equipped with a heated (H) electrospray ionization (ESI) interface (Thermo Fisher Scientific, Waltham, Massachusetts). Two analytical sequences (one with positive and one with negative ionization mode) were executed with a “Full Scan-data dependent fragmentation” experiment.

Source conditions were as follows: sheath and auxiliary gas flow rates of 40 and 20 arbitrary units, respectively; heater temperature of 250°C with a spray voltage of 3.2 kV (ESI pos) and −3.0 kV (ESI neg). The capillary was kept at 220°C and the S-lens RF level was set at 55 AU for both the acquisition modes.

The full-scan accurate mass spectra from 75 to 1000 Da for both the ionization modes were obtained with a resolution of 70 000 Full Width at Half Maximum (FWHM) (m/z 200), automatic gain control target 1e6, and maximum injection time of 200 milliseconds.

In all the experiments, the data-dependent tandem mass spectrometry (MS/MS) acquisition was executed with a resolution of 17 500 FWHM (m/z 200) and intensity threshold 6e4. The quadrupole isolation window was kept at 2.0 m/z with a TopN value of 5. The scan range was from 50 to the fragmented mass m/z (m/z +25), automatic gain control target 2e5, maximum injection time of 50 milliseconds, and normalized collision energy of 30% with ±50% step.

Samples, together with the “extraction blanks,” were randomly injected to avoid systematic bias.16 The “standard solutions” and the QC samples were injected at the beginning of the sequences and every 10 sample injections.

### Table 1: Sampling plan for the model creation

| Sample | 0 h | 1 d | 2 d |
|--------|-----|-----|-----|
| 1      | X   | X   | X   |
| 2      | X   | X   | X   |
| 3      | X   | X   | X   |
| 4      | X   | X   | X   |
| 5      | X   | X   | X   |
| 6      | X   | X   | X   |
| 7      | X   | X   | X   |
| 8      | X   | X   | X   |
| 9      | X   | X   | X   |
| 10     | X   | X   | X   |
| 11     | X   | X   | X   |
| 12     | X   | X   | X   |
| 13     | X   | X   | X   |
| 14     | X   | X   | X   |
| 15     | X   | X   | X   |
| 16     | X   | X   | X   |
| 17     | X   | X   | X   |
| 18     | X   | X   | X   |
| 19     | X   | X   | X   |
| 20     | X   | X   | X   |
| 21     | X   | X   | X   |
| 22     | X   | X   | X   |
| 23     | X   | X   | X   |
| 24     | X   | X   | X   |
| 25     | X   | X   | X   |
| 26     | X   | X   | X   |
| 27     | X   | X   | X   |
| 28     | X   | X   | X   |
| 29     | X   | X   | X   |

*aDouble sample prep.*

For the evaluation of method reproducibility, 20% of the samples were double prepared, as detailed in Table 1.

During each extraction session, the same procedure was performed also into empty tubes, in which all the steps were executed without the egg product addition. These samples were labelled as “extraction blanks.”

A 500-ng/mL chloramphenicol solution in acetonitrile to water ratio of 80:20 was prepared and named “standard solution.”

Additionally, two quality control (QC) samples were prepared mixing 10 μL of each extract sample.

### 2.4 LC-HRMS analysis

High-performance liquid chromatography analysis was performed with a Dionex UltiMate 3000 ultrahigh-performance liquid chromatography (Thermo Fisher Scientific, Inc, Waltham, Massachusetts) equipped with a Luna Omega 150 mm × 2.1 mm, 1.6 μm particle size analytical column (Phenomenex, Torrance, California) with a precolumn security guard ultra C18, both maintained at 30°C. Gradient elution was performed using FA and AF as mobile phase modifiers with a constant flow rate of 0.3 mL/min.

Ultra-high-performance liquid chromatography high-resolution mass spectrometry raw data were acquired using Xcalibur software (version 3.0, Thermo Fisher Scientific, Waltham, Massachusetts); peaks alignment, “extraction blanks” subtraction, and features extraction were performed using Compound Discoverer software (version 2.0 Thermo Fisher Scientific, Waltham, Massachusetts); the mass range inspected was between 50 m/z and 1000 m/z from 1 to 30 minutes of the chromatographic runs.

The values of the critical parameters for features extractions are the following: precursor ion deviation of 5 ppm (for the positive runs) and 10 ppm (for the negative runs); maximum retention time shift of 0.3 minutes; minimum peak intensity for a peak to be selected 1 000 000 AU; relative intensity tolerance used for isotope search 30%.

Structures prediction was also performed using “ChemSpider” databases, setting a maximum mass shift of 5 ppm for positive acquisitions and 10 ppm for negative acquisitions.

For the “m/z CLOUD” MS/MS library search, the precursor mass tolerance used was 0.05 Da while the fragments mass tolerance was 10 ppm.
The resulting two data matrices (for positive and negative ionization modes), containing the area values provided by Compound Discoverer for all the features, were exported and processed with SIMCA software (version 14.1; Umetrics, Umea, Sweden) for chemometric data elaboration.

Data were log transformed and Pareto scaled, and then a preliminary principal component analysis (PCA) was executed to check the clustering of the samples and the QCs positioning in the scores plot.

Subsequently, features were filtered, and only the ones that had a Coefficient of Variation (CV)% lower than 40% in the QC samples were selected.

A new PCA was executed to check the expected improvement in samples separation and QCs positioning; after the evaluation of the replicate samples placement in the scores plot, a final PCA was calculated without the QC samples and with the average values of the replicates.

Afterwards, supervised orthogonal partial least square discriminant analysis (OPLS-DA) models were built comparing the “fresh” samples against the “1 day” samples, against the “2 days” samples and against the union of “1 day” and “2 days” samples.

Thanks to the S-plots, statistically significant markers responsible of the clustering were selected: The ions furthest away from the origin contribute more significantly to the separation between the groups and may therefore be regarded as the differentiating ones. This approach was used for the six OPLS-DA models to be sure that all the most discriminative features could be selected.

In addition, their VIP (variable influence on projection) values were evaluated to assess their relevance for the chemometric model (VIP values had to be >1.4). Features that survived this process were studied, and a tentative of compounds identification was performed.

In all the chemometric models created, an internal leave 1/7 out cross-validation was executed.

### 2.6 Compounds identification workflow

Compounds identification was performed according to the following steps:

1. Chemical formula hypothesis, taking into account also adducts, if present.
2. Retention time evaluation.
3. Evaluation of the mass shift between the experimental and the theoretical values.
4. Isotopic pattern evaluation and hypothesis of a chemical structure referring also to ChemSpider databases.
5. MS/MS spectra study and comparison between the theoretical fragmentations (performed with Compound Discoverer) and the fragments detected.
6. MS/MS spectra comparison with the m/z CLOUD database (in which fragmentations with other Orbitrap instruments are available).
7. If available, Reference Standard injection as final confirmation.

The Standard Initiative in Metabolomics and subsequent improvements suggest different ranks of compounds identification according to the number of steps completed: identification steps completed: 1 to 5, level of identification: 3; identification steps completed: 1 to 6, level of identification: 2; identification steps completed: 1 to 7, level of identification: 1; this classification was used for the selected features.

### 2.7 Markers validation procedure

The compounds identified with the workflow detailed above were searched in seven new egg products batches not used for the model creation; they were collected, extracted, and analyzed with the same experimental design previously described.

The maximum reliability of these compounds as “freshness” or “not freshness” markers implies they have to be confirmed not only with their presence or absence in the new samples but also highlighting the same increasing or decreasing trend through the time points.

### 3 RESULTS AND DISCUSSION

#### 3.1 Extraction procedures and sequences evaluation

Before starting the real data elaboration, different evaluations on the internal standard results were performed.

Chloramphenicol is an exogenous compound that was added to every sample with the aim of monitoring the extraction procedures and the goodness of the analytical sequences.

For the sequence analyzed in positive ionization mode, the values obtained were 10.8% (areas) and 2.2% (retention time) and, for the negative ionization mode, 10.1% (areas) and 2.8% (retention time).

The results obtained highlight that the response of the instrument was acceptable during the whole sequences (126 injections each) for both positive and negative ionization modes.

An analogous evaluation of chloramphenicol peaks intensity and retention times shift was performed for all the sample injections: In this way, more information about the goodness of the analytical sequences was obtained; moreover, the CV% of the area values of the internal standard was used for the extraction reproducibility evaluation, assuring that no mistakes occurred during samples preparation.

For the sequence analyzed in positive ionization mode, the areas CV was 7.2%, and the retention time CV was 1.6% and, for the negative ionization mode, 10.5% (areas) and 3.1% (retention time).

According to the results obtained, we can deduce that no issues occurred during sample extraction procedures; moreover, the slight shift in the retention times is a further proof of the goodness of the analytical sequences.

#### 3.2 Chemometric data elaboration

The preliminary PCA obtained with the areas of the entire data set highlighted a separation between the fresh samples and the not fresh
ones with a clustering of the QC samples for both positive and negative ionization modes (data not shown); the first two components described more than 45% of the variance for both the models.

After data filtering (obtained by selecting in the QC samples the features with CV lower than 40% in the area values), an improvement in the clusterization of the samples was obtained, with a tight clustering of the QC samples (Figures 1 and 2). In addition, the overlay in the scores plot of the replicate extracted samples was essentially complete.

The tight clustering of the QC samples in the center of the score plot is a further confirmation of the robustness of the analytical procedure; moreover, this ensure that the separation through the groups is not random but due to a real variability.26

The overlay in the scores plot of the replicate samples certify that the extraction procedure is repeatable not only for the chloramphenicol compound but for the whole fingerprint of the egg product.

The final PCA, created with the average value of the replicates and without the QC samples, clearly highlights the separation between the fresh samples and the not fresh ones (Figures 3 and 4).

All the OPLS-DA–supervised models, as expected, extraordinarily increase the separation between the two groups. As example, Fig-

FIGURE 1  ESI + PCA scores plot of the samples (n = 107) after features filtering. Leftmost area dots (0 h), fresh samples; 1D dots, “1 day” samples; 2D dots, “2 days” samples; Quality Control (QC) dots, “QC samples” (circled). Explained variance of the first two PCs, 55.6%. ESI, electrospray ionization; PCA, principal component analysis; QC, quality control

FIGURE 2  ESI – PCA scores plot of the samples (n = 107) after features filtering. Leftmost area dots (0 h), fresh samples; 1D dots, “1 day” samples; 2D dots, “2 days” samples; QC dots, “QC samples” (circled). Explained variance of the first two PCs, 55.9%. ESI, electrospray ionization; PCA, principal component analysis; QC, quality control

FIGURE 3  Final ESI + PCA scores plot of the samples. Leftmost area dots (0 h), fresh samples; 1D dots, “1 day” samples; 2D dots, “2 days” samples. Explained variance of the first two PCs, 57.3%. ESI, electrospray ionization; PCA, principal component analysis; QC, quality control
FIGURE 4 Final ESI – PCA scores plot of the samples. Leftmost area dots (0 h), fresh samples; 1D dots, “1 day” samples; 2D dots, “2 days” samples. Explained variance of the first two PCs, 58.2%. ESI, electrospray ionization; PCA, principal component analysis; QC, quality control [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 5 ESI + OPLS-DA scores plot of the fresh samples against the “1 day” samples. Left area dots (0 h), fresh samples; right area dots (1D), “1 day” samples. ESI, electrospray ionization; OPLS-DA, orthogonal partial least square discriminant analysis [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 6 ESI – OPLS-DA scores plot of the fresh samples against the “1 day” samples. Left area dots (0 h), fresh samples; right area dots (1D), “1 day” samples. ESI, electrospray ionization; OPLS-DA, orthogonal partial least square discriminant analysis [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Global value of R²X (cum), R²Y (cum), and Q² parameters for positive ionization mode

| Model     | PCA (Preliminary) | PCA (Filtered) | PCA (Final) | OPLS-DA ("Fresh" vs "1D + 2D") | OPLS-DA ("Fresh" vs "1D") | OPLS-DA ("Fresh" vs "2D") |
|-----------|-------------------|----------------|-------------|---------------------------------|---------------------------|---------------------------|
| R²X (cum) | 0.691             | 0.750          | 0.725       | 0.618                           | 0.589                     | 0.680                     |
| R²Y (cum) | /                 | /              | /           | 0.992                           | 0.997                     | 0.999                     |
| Q²        | 0.544             | 0.643          | 0.590       | 0.989                           | 0.990                     | 0.994                     |

Abbreviations: OPLS-DA, orthogonal partial least square discriminant analysis; PCA, principal component analysis.
/ means data available only in OPLS-DA mode and not in PCA mode.

Tables 2 and 3 summarize the variance of the x and y variables explained by the models previously described—R²X (cum) and R²Y.
The results obtained strongly indicate that the models created for both positive and negative ionization modes are reliable and robust.

### 3.3 Compounds identification

Features responsible of the clusterization were selected, and a group of them was identified according to what described in the “compounds identification workflow” paragraph.

Both markers of “not freshness” (that drastically increase their intensity or even appear during the eggs ageing) and of “freshness” (that drastically decrease their intensity or even disappear during the eggs ageing) were identified.

Tables 4 and 5 summarize the compounds responsible of the “freshness” and “not freshness” classification, ranked according to the level of identification suggested by the Standard Initiative in Metabolomics.

### 3.4 Markers validation

As also suggested by the US Pharmacopoeia, predictive chemometric models should be always validated with an external set of samples (possibly collected in different periods) that have to be treated as unknown; their classification should be predicted, to certify that the

---

### TABLE 3 Global value of $R^2_X$ (cum), $R^2_Y$ (cum), and $Q^2$ parameters for negative ionization mode

| Model       | PCA (Preliminary) | PCA (Filtered) | PCA (Final) | OPLS-DA (“Fresh” vs “1D + 2D”) | OPLS-DA (“Fresh” vs “1D”) | OPLS-DA (“Fresh” vs “2D”) |
|-------------|-------------------|----------------|-------------|--------------------------------|---------------------------|---------------------------|
| $R^2_X$ (cum) | 0.750             | 0.769          | 0.777       | 0.617                          | 0.627                      | 0.698                      |
| $R^2_Y$ (cum) | /                 | /              | /           | 0.994                          | 0.997                      | 0.999                      |
| $Q^2$        | 0.584             | 0.577          | 0.546       | 0.991                          | 0.991                      | 0.996                      |

Abbreviations: OPLS-DA, orthogonal partial least square discriminant analysis; PCA, principal component analysis.

---

### TABLE 4 Resume of “not freshness” markers identified

| Name                        | Pseudomolecular Ion         | Detected $m/z$ | RT, min | Predicted Formula | Mass Error, ppm | CV IN QCs | VIP Value | ID Type |
|-----------------------------|-----------------------------|----------------|---------|-------------------|-----------------|-----------|-----------|---------|
| Lactic acid                 | [M – H]^-                    | 89.0244        | 1.73    | C3 H6 O3          | 12.4^d          | 27        | 2.59^c    | 1       |
| Succinic acid               | [M – H]^+                    | 165.0555       | 11.95   | C9 H10 O3         | 5.45            | 11        | 2.64^b    | 1       |
| Tyramine                    | [M + H]^+                    | 117.0193       | 3.22    | C4 H6 O4          | 8.55            | 11        | 2.49^b    | 1       |
| [M + H-NH3]^+               | 138.0913                    | 3.65            | C8H11NO | 0.72              | 30              | 2.21^b    | 1       |
| [M + H]^-                   | 111.0196                    | 1.65            |         |                   |                 |           |           |         |
| 4-Hydroxybutyric acid       | [M – H]^-                    | 103.0398       | 3.90    | C4 H8 O3          | 8.74            | 23        | 1.83^b    | 2       |
| 6-Methylquinoline           | [M + H]^+                    | 144.0807       | 8.39    | C10 H9 N          | 0.69            | 7         | 2.27^b    | 2       |
| L-Alanine-D-alanine         | [M – H]^+                    | 150.0773       | 1.34    | C6 H12 N2 O3      | 5.66            | 12        | 1.88^b    | 2       |
| Hydroxycaproic acid         | [M – H]^-                    | 131.0712       | 11.25   | C6 H12 O3         | 7.63            | 17        | 3.57^b    | 2       |
| Phosphorylethanolamine      | [M – H]^-                    | 140.0118       | 1.14    | C2 H8 N O4 P      | 7.85            | 18        | 3.00^b    | 2       |
| 3-Thiomorpholinecarboxylic acid | [M + H]^+                  | 148.0426       | 1.70    | C5 H9 N O2 S      | 0.68            | 29        | 2.09^b    | 3       |
| Arginine-proline dipeptide  | [M + H]^+                    | 272.1716       | 1.30    | C11 H21 N5 O3     | 0.37            | 13        | 2.55^b    | 3       |
| Lysine leucine dipeptide    | [M – H]^-                    | 258.1818       | 1.70    | C12 H25 N3 O3     | 2.32            | 16        | 1.81^b    | 3       |
| N-Acetylhistidine           | [M – H]^-                    | 196.0725       | 1.43    | C8 H11 N3 O3      | 4.08            | 17        | 2.20^b    | 3       |
| 2-Hydroxyisovaleric acid or 2-hydroxymethylbutyric acid | [M – H]^- | 117.0557 | 7.15 | C5 H10 O3 | 9.39 | 18 | 2.15^b | 3 |
| Mannose 6-phosphate or glucose 6-phosphate | [M – H]^- | 259.0226 | 1.08 | C6 H13 O9 P | 5.01 | 32 | 2.99^c | 3 |

Abbreviations: OPLS-DA, orthogonal partial least square discriminant analysis; QC, quality control; RT, retention time; VIP, variable influence on projection.

^aVIP value obtained from the OPLS-DA model 0 h vs all.
^bVIP value obtained from the OPLS-DA model 0 h vs 2D.
^cVIP value obtained from the OPLS-DA model 0 h vs 1D.
^dThis value is higher than the limit (10 ppm) but is due to the low molecular weight of lactic acid. The injection of the reference standard certified its identity beyond any reasonable doubt.
samples clustering is real and not related to some overfitting of the
chemometric model.

This approach is hard to perform with metabolomic results: Because of the ultrahigh-performance liquid chromatography–high-resolution mass spectrometry intrinsic variability, only chromatograms acquired with the same analytical sequence can be aligned and studied; the use of chemometric model to predict the classification of samples acquired with different sequences could lead to dangerous mistakes. Moreover, the goal of this work was the identification of new markers of freshness, and the chemometric model was considered only a tool to reach this goal.

For these reasons, the marker compounds and not the model were validated, that means that all these 31 features were searched in the new samples and their presence or absence, together with their trends through the time points, were evaluated and compared with the results obtained during the model creation, as presented in Tables 7 and 8.

All the target molecules were detected with almost the same trend through the time points (with exception of a couple of markers but only in terms of relative ratio): These results prove that the compounds identified are real markers of “freshness” or “not freshness” in egg products and are not related only to a misleading overfitting of the chemometric model.

### 3.5 Markers interpretive

Most of the markers reported in this work have not been previously identified in eggs shelf life studies, because of the analytical approach. While egg freshness is commonly followed by volatile profiling or by e-nose fingerprinting, the use of high-resolution liquid chromatography–mass spectrometry allowed to enlarge the identification to medium-polar nonvolatile compounds.

Volatile markers are actually based on the formation of volatile aldehydes from fatty acids oxidative degradation. All the target molecules were detected with almost the same trend through the time points (with exception of a couple of markers but only in terms of relative ratio): These results prove that the compounds identified are real markers of “freshness” or “not freshness” in egg products and are not related only to a misleading overfitting of the chemometric model.

### TABLE 5 Resume of “freshness” markers identified

| Name                                      | Pseudomolecular Ion | Detected m/z | RT, min | Predicted Formula | Mass Error, ppm | CV in QCs, % | VIP Value | ID Type |
|-------------------------------------------|---------------------|--------------|---------|-------------------|-----------------|--------------|-----------|---------|
| Asparagine                                | [M + H]^+           | 133.0608     | 1.21    | C4 H8 N2 O3       | 0.30            | 14           | 2.35^c    | 1       |
| Glutamine                                 | [M + H]^+           | 147.0762     | 1.23    | C5 H10 N2 O3      | 1.36            | 8            | 2.75^c    | 1       |
|                                           | [M – H]^-           | 145.0615     | 1.17    |                   | 4.83            | 12           | 2.55^c    | 1       |
| Guanosine                                 | [M – H]^-           | 282.0844     | 4.85    | C10 H13 N5 O5     | 3.90            | 33           | 2.88^c    | 1       |
| Guanosine monophosphate (GMP)             | [M – H]^-           | 362.0506     | 1.66    | C8 H19 N3 O9 P2   | 1.93            | 28           | 3.10^c    | 1       |
| Serine                                    | [M + H]^+           | 106.0502     | 1.23    | C3 H7 N O3        | 2.83            | 8            | 2.67^c    | 1       |
| Uridine                                   | [M – H]^-           | 243.0619     | 1.68    | C9 H12 N2 O6      | 2.88            | 16           | 2.85^c    | 1       |
| Uridine monophosphate                     | [M – H]^-           | 323.0285     | 1.37    | C9 H13 N2 O9 P   | 3.10            | 17           | 3.20^c    | 1       |
| Arginine                                  | [M + H]^+           | 175.1189     | 1.22    | C6 H14 N4 O2      | 0.22            | 16           | 2.15^b    | 2       |
|                                           | [M – H]^-           | 173.1038     | 1.17    | C6 H9 N3 O2       | 2.89            | 9            | 1.69^b    | 2       |
| Aspartic acid                             | [M + H]^+           | 134.0448     | 1.22    | C4 H7 N O4        | 0.03            | 10           | 1.72^a    | 2       |
| Histidine                                 | [M + H]^+           | 156.0768     | 1.20    | C6 H9 N3 O2       | 0.02            | 21           | 1.69^b    | 2       |
| Methionine sulfoxide                      | [M + H]^+           | 166.0530     | 1.28    | C5 H11 N O3 S    | 1.20            | 12           | 1.68^b    | 2       |
| Methylhistidine                           | [M + H]^+           | 170.0924     | 1.22    | C7 H11 N3 O2      | 0.06            | 37           | 2.50^c    | 2       |
| Threonine                                 | [M + H]^+           | 120.0656     | 1.25    | C4 H9 N O3        | 0.80            | 8            | 1.42^a    | 2       |
| Guanosine 5′′-diphospho-D-mannose or Guanosine 5′′-diphospho-D-glucose | [M – H]^- | 604.0694 | 1.30 | C29 H17 N7 O5 P2 | 1.99 | 27 | 2.77^c | 3 |
| N-Acetyl-α-D-galactosamine or N-acetyl-α-D-glucosamine | [M + H]^+ | 222.0973 | 1.38 | C8 H15 N O6 | 0.45 | 17 | 1.87^a | 3 |

Abbreviations: OPLS-DA, orthogonal partial least square discriminant analysis; QC, quality control; RT, retention time; VIP, variable influence on projection.

^aVIP value obtained from the OPLS-DA model 0 h vs all.

^bVIP value obtained from the OPLS-DA model 0 h vs 2D.

^cVIP value obtained from the OPLS-DA model 0 h vs 1D.

### TABLE 6 Resume of the features selected after each statistical step

| Data Analysis Step                      | Number of Features Selected |
|-----------------------------------------|-----------------------------|
| Peak alignment                          | 3452                        |
| Filtration according to CV% in QC       | 2189                        |
| Extracted features from S-plots         | 207                         |
| Final list                              | 14                          |

Abbreviation: QC, quality control.
| Compound Name                                         | Model Creation | Marker Validation |
|-------------------------------------------------------|----------------|------------------|
| Lactic acid                                           | 1E+08          | 6E+07            |
| Phenyllactic acid                                     | 1E+07          | 5E+06            |
| Succinic acid                                         | 8E+06          | 2E+07            |
| Tyramine                                              | 2E+07          | 5E+06            |
| Uracil                                                | 9E+05          | 5E+05            |
| 2-Hydroxyisovaleric acid or 2-hydroxymethylbutyric acid | 6E+06          | 1E+06            |
| 4-Hydroxybutyric acid                                 | 7E+05          | 3E+05            |
| 6-Methylquinoline                                     | 2E+07          | 1E+07            |
| D-Alanyl-D-alanine                                    | 3E+05          | 2E+05            |
| Hydroxycaproic acid                                   |                |                  |

(Continues)
### TABLE 7  (Continued)

| Compound Name                              | Model Creation | Marker Validation |
|--------------------------------------------|----------------|-------------------|
| Mannose 6-phosphate or glucose 6-phosphate | ![Graph](image1) | ![Graph](image2) |
| Phosphorylethanolamine                     | ![Graph](image3) | ![Graph](image4) |
| 3-Thiomorpholinecarboxylic acid           | ![Graph](image5) | ![Graph](image6) |
| Arginine-proline dipeptide                | ![Graph](image7) | ![Graph](image8) |
| Lysine leucine dipeptide                  | ![Graph](image9) | ![Graph](image10) |
| N-Acetylhistidine                         | ![Graph](image11) | ![Graph](image12) |

*Time points: first bar, "t zero"; second bar, "1 day"; third bar, "2 days."

### TABLE 8  Comparison of the mean area values (+/− standard error) of freshness markers between the model creation and the markers validation*

| Compound Name                              | Model Creation | Marker Validation |
|--------------------------------------------|----------------|-------------------|
| Asparagine                                 | ![Graph](image13) | ![Graph](image14) |
| Glutamine (ESI+)                           | ![Graph](image15) | ![Graph](image16) |

(Continues)
| Compound Name                | Model Creation | Marker Validation |
|-----------------------------|----------------|-------------------|
| Glutamine (ESI–)            |                |                   |
| Guanosine                   |                |                   |
| Guanosine monophosphate     |                |                   |
| Serine                      |                |                   |
| Uridine                     |                |                   |
| Uridine monophosphate       |                |                   |
| Arginine (ESI+)             |                |                   |
| Arginine (ESI–)             |                |                   |
| Aspartic acid               |                |                   |
markers deriving from the very first stage of fatty acid oxidative degradation are reported in the fresh products (namely, 4-hydroxybutyric acid, hydroxycaproic acid, 2-hydroxyvaleric acid). These are probably formed during thermal stabilization of egg products, accumulated over time, and further degraded to volatile aldehydes along storage.

Interestingly, most of the compounds reported as markers of freshness are precursors of compounds listed as “markers of nonfreshness.” In particular, it can be observed a strong activation of nitrogen and pyrimidine metabolism pathways, probably on account of the metabolic activity of residual microorganisms growing over time. This leads to the decrease of free amino acids, with the accumulation of biogenic amine (ie, tyramine), amino acid, and purine degradation products (ie, phosphorylethanolamine, N-acetylhistidine, phenyllactic acid, and uracil), and precursors of peptidoglycane (D-alanylglycanal). Consistently, an activation of microbial fermentation is also attested by the accumulation of succinic and lactic acid over time.

### 4 CONCLUSIONS

In the current study, new compounds related to freshness in egg products samples were identified using a nontargeted metabolomic
approach. After model creation, the robustness of the identified markers of freshness was assessed through the analysis of a validation set of not previously used egg products.

Along with lactic acid and succinic acid, the reference compounds already considered in the EU legislation, other 29 compounds were detected as discriminant features, allowing a more robust evaluation of freshness.

Further improvements of the results presented in this study should lead to two directions. Firstly, target methods on the identified compounds could be developed for a quantitative evaluation: Quality control laboratories should only be able to detect these 31 molecules, avoiding the need of high-resolution mass spectrometry and chemometric software, that are much more expensive than a simple single-stage liquid chromatography–mass spectrometry instrument.

Secondly, this group of compounds, and generally the recorded fingerprints, could be helpful also for both the safety evaluations on microbial growth perspectives and the detection of other frauds related to egg products food chain, as, for example, the illegal use of incubated eggs.

ACKNOWLEDGEMENTS

Authors would like to thank Angelo Restori and all the Barilla’s quality control laboratory team for the invaluable support.

This work was supported by the European Union, Seventh Framework Program (FP7/2007-2013), under grant agreement 613688 (Ensuring the Integrity of the European food chain FoodIntegrity).

ORCID

Michele Suman @ http://orcid.org/0000-0003-0719-6627

REFERENCES

1. D.L. no. 65, February 1993 Italian Legislation reception of European Union - https://www.tuttocamere.it/files/attivita/1993_65.pdf
2. European Regulation Reg. CE 853/2004 - http://eur-lex.europa.eu/legal-content/IT/TXT/?uri=LEGISSUM%3Af84002
3. Li J, Zhu S, Jiang S, Wang J. Prediction of egg storage time and yolk index based on electronic nose combined with chemometric methods. LWT - Food Sci Technol. 2017;82:369-376.
4. Suman M, Riani G, Dalcanele E. MOS-based artificial olfactory system for the assessment of egg products freshness. Sens Actuators B, Chemical. 2007;125(1):40-47.
5. Yimenu S, Kim J, Kim B. Prediction of egg freshness during storage using electronic nose. Poult Sci. 2017;96(10):3733-3746.
6. Wang Q, Zhou K, Wang C, Ma M. Egg freshness detection based on hyperspectral image technology. Adv J Food Sci Technol. 2015;7(8): 652-657.
7. Zhao J, Li H, Chen Q, Huang X, Sun Z, Zhou F. Identification of egg’s freshness using NIR and support vector data description. J Food Eng. 2010;98(4):408-414.
8. Lin H, Zhao J, Sun L, Chen Q, Zhou F. Freshness measurement of eggs using near infrared (NIR) spectroscopy and multivariate data analysis. Innov Food Sci Emerg Technol. 2011;12(2):182-186.
9. Riedl J, Esslinger S, Fauli-Hassek C. Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. Anal Chim Acta. 2015;885:17-32.
10. Figueira J, Cámara H, Pereira J, Câmara J. Evaluation of volatile metabolites as markers in Lycopersicon esculentum L. cultivars discrimination by multivariate analysis of headspace solid phase microextraction and mass spectrometry data. Food Chem. 2014;145:653-663.
11. Silva P, Freitas J, Silva C, Perestrello R, Nunes F, Câmara J. Establishment of authenticity and typicality of sugarcane honey based on volatile profile and multivariate analysis. Food Control. 2017;73:1176-1188.
12. Aliakbarzadeh G, Sereshht H, Parastar H. Pattern recognition analysis of chromatographic fingerprints of Crocus sativus L. secondary metabolites towards source identification and quality control. Anal Bioanal Chem. 2016;408(12):3295-3307.
13. Hajslova J, Cajka T, Vlaclavik L. Challenging applications offered by direct analysis in real time (DART) in food-quality and safety analysis. Trends Anal Chem. 2011;30(2):204-218.
14. Balog J, Perenyi D, Guillon-Hoyas C, et al. Identification of the species of origin for meat products by rapid evaporative ionization mass spectrometry. J Agric Food Chem. 2016;64(23):4793-4800.
15. Jandric Z, Islam M, Singh D, Cannavan A. Authentication of Indian citrus fruit/fruit juices by untargeted and targeted metabolomics. Food Control. 2017;72:181-188.
16. Righetti L, Rubert J, Galaverna G, et al. A novel approach based on untargeted lipidomics reveals differences in the lipid pattern among durum and common wheat. Food Chem. 2018;240:775-783.
17. Kalogouri N, Alygizakis N, Aalizadeh R, Thomaidis N. Olive oil authenticity studies by target and nontarget LC–QTOF-MS combined with advanced chemometric techniques. Anal Bioanal Chem. 2016;408(28):7955-7970.
18. Collins T, Zweigenbaum J, Ebeler S. Profiling of nonvolatiles in whiskies using ultra high pressure liquid chromatography quadrupole time-of-flight mass spectrometry (UHPLC-QTOF MS). Food Chem. 2014;163:186-196.
19. Rubert J, Zachariasova M, Hajslova J. Advances in high-resolution mass spectrometry based on metabolomics studies for food— a review. Food Addit Contam Part a. 2015;32(10):1685-1708.
20. US Pharmacopeial Convention. Guidance on developing and validating nontargeted methods for adulteration detection. 2016, Appendix XVIII, 2053.
21. Jandric Z, Roberta D, Rathor M, Abraham A, Islam M, Cannavan A. Assessment of fruit juice authenticity using UPLC–QTOF MS: a metabolomics approach. Food Chem. 2014;148:7-17.
22. Akarachantachote N, Chadcham S, Saithanu K. Cutoff threshold of variable importance in projection for variable selection. Int J Pure Appl Math. 2014;94(3):307.
23. Sumner L, Amberg A, Barrett D, et al. Proposed minimum reporting standards for chemical analysis. Metabolomics. 2007;3(3):211-221.
24. Schymanski E, Jeon J, Gulde R, et al. Identifying small molecules via high resolution mass spectrometry data. Environ Sci Technol. 2014;48(4):2097-2098.
25. Khamis M, Adamko D, El-Meuther A. Mass spectrometry based approaches in urine metabolomics and biomarker discovery. Mass Spectrom Rev. 2017;36(2):115-134.
26. Godzien J, Alonso-Herranz V, Barbas C, Armitage E. Controlling the quality of metabolomics data: new strategies to get the best out of the QC sample. Metabolomics. 2015;11(3):518-528.

How to cite this article: Cavanna D, Catellani D, Dall’Ast C, Suman M. Egg product freshness evaluation: A metabolomic approach. J Mass Spectrom. 2018;53:849–861. https://doi.org/10.1002/jms.4256