

This is the peer reviewed version of the following article:

The potential of spectral and hyperspectral-imaging techniques for bacterial detection in food: A case study on lactic acid bacteria / Foca, Giorgia; Ferrari, Carlotta; Ulrici, Alessandro; Sciutto, Giorgia; Prati, Silvia; Morandi, Stefano; Brasca, Milena; Lavermicocca, Paola; Lanteri, Silvia; Oliveri, Paolo. - In: TALANTA. - ISSN 0039-9140. - 153:(2016), pp. 111-119. [10.1016/j.talanta.2016.02.059]

Terms of use:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

19/12/2025 19:37

The potential of spectral and hyperspectral-imaging techniques for bacterial detection in food: a case study on lactic acid bacteria

Giorgia Foca^a, Carlotta Ferrari^a, Alessandro Ulrici^a, Giorgia Sciutto^b, Silvia Prati^b, Stefano Morandi^c, Milena Brasca^c, Paola Lavermicocca^d, Silvia Lanteri^e, Paolo Oliveri^{e,*}

^a*University of Modena and Reggio Emilia, Department of Life Sciences, Via Amendola 2, 42122 Reggio Emilia, Italy*

^b*University of Bologna, Ravenna Campus, Microchemistry and Microscopy Art Diagnostic Laboratory (M2ADL), Via Guaccimanni 42, 48100 Ravenna, Italy*

^c*National Research Council, Institute of Sciences of Food Production (CNR-ISPA), Via Celoria 2, 20133 Milano, Italy*

^d*National Research Council, Institute of Sciences of Food Production (CNR-ISPA), Via Amendola 122/O, 70126 Bari, Italy*

^e*University of Genova, Department of Pharmacy, Via Brigata Salerno 13, 16147 Genova, Italy*

*Corresponding author:

Paolo Oliveri Ph.D.

University of Genoa – Dept. of Pharmacy

Via Brigata Salerno, 13

I-16147 GENOA

Tel: +39 010 3532626

Fax: +39 010 3532684

e-mail: oliveri@dictfa.unige.it

48 **Abstract**

49 Official methods for the detection of bacteria are based on culture techniques. These methods have
50 limitations such as time consumption, cost, detection limits and the impossibility to analyse a large
51 number of samples. For these reasons, the development of rapid, low-cost and non-destructive
52 analytical methods is a task of growing interest.

53 In the present study, the capability of spectral and hyperspectral techniques to detect bacterial
54 surface contamination was investigated preliminarily on gel cultures, and subsequently on sliced
55 cooked ham. In more detail, two species of lactic acid bacteria (LAB) were considered, namely
56 *Lactobacillus curvatus* and *Lactobacillus sakei*, both of which are responsible for common
57 alterations in sliced cooked ham.

58 Three techniques were investigated, with a different equipment, respectively: a macroscopic
59 hyperspectral scanner operating in the NIR (10,470-5880 cm^{-1}) region, a FT-NIR
60 spectrophotometer equipped with transmission arm as sampling tool, working in the 12,500-5800
61 cm^{-1} region, and a FT-MIR microscopy operating in the 4000-675 cm^{-1} region.

62 Multivariate exploratory data analysis, in particular principal component analysis (PCA), was
63 applied in order to extract useful information from original data and from hyperspectrograms. The
64 results obtained demonstrate that the spectroscopic and imaging techniques investigated can
65 represent an effective and sensitive tool to detect surface bacterial contamination in samples and, in
66 particular, to recognise the species to which the bacteria belong.

67

68 **Keywords**

69 FT-NIR spectroscopy; FT-IR microscopy; hyperspectral imaging; principal component analysis
70 (PCA); lactic acid bacteria (LAB); cooked ham.

71

72

73 **Highlights**

- 74 ➤ Official methods for detecting bacteria are culture-dependent
75 ➤ Limitations of these methods are high time consumption, costs and detection limits
76 ➤ Capability of sensitive spectral and hyperspectral techniques was investigated
77 ➤ The potential of near and the infrared spectral regions were explored
78 ➤ Multivariate exploratory data analysis was applied to extract useful information

79

80

81

1. Introduction

Food labelling regulations highlight the importance of providing the consumer with exact information about ingredients and additives present in food [1,2]. Such a key concern is reported also in the recent UE regulations, which applies from 2014 [3]. In particular, commission regulations also lay down microbiological criteria for foodstuffs, maximum limits for bacterial contamination, toxins and biogenic amines [4]. A direct consequence of this is the requirement for new analytical methodologies aimed at simplifying the use and improving the efficiency of existing control tools.

The shelf-life of cooked and sliced meat products, like cooked ham, is limited mainly because of microbiological safety and spoilage issues. This is because manipulations like slicing and packaging unavoidably reintroduce bacterial contaminants after the cooking process, and because the product has a near-neutral pH (around 6) and water activity higher than 0.945 [5,6].

Spoilage of packaged sliced ham is mostly accompanied by souring, slimy meat juice exudates and swelling of the pack due to gas production and is usually caused by lactic acid bacteria (LAB), together with *Pseudomonas*. Spoilage of cooked meat products results in sensory quality defects such as sour off-flavour, discolouration, gas production, and ropy slime formation.

Official methods for the detection of these bacterial species (ISO 15214:1998, ISO 13720:2010) are based on culture techniques. These methods have limitations such as time consumption, cost, detection limits and the impossibility to analyse a large number of samples. For these reasons, the development of rapid, low-cost and non-destructive analytical methods is a task of growing interest. Methods based on classical spectroscopy and on the analysis of conventional images (RGB) have been studied extensively for several years [7]. Although applied to numerous food issues, these conventional approaches have a limited capacity for obtaining information from food samples, discarding either the spatial information – in the former approach – or the spectral one – in the latter [8]. In recent years, hyperspectral imaging (HSI) has emerged as a very efficient solution for the quality and safety control of food products, being able to thoroughly characterise individual components within complex matrices, allowing both chemical identification and localisation [9–13]. In fact, if the conventional image analysis gives an answer to the question 'where' and conventional spectroscopy may give answers to the two questions 'what' and 'how much', HSI gives an answer to the combined question 'where and how much of what', thus providing a comprehensive characterisation, particularly useful for identification and control purposes. In addition, collecting spectral profiles at multiple points in a given area increases the representativeness of the information obtained.

115 HSI can be used either on a macroscopic or a microscopic scale and in different spectral ranges,
116 depending on the level of investigation and the spatial resolution required, the chemical information
117 of interest to the particular issue and the type of implementation needed. The mid and near infrared
118 regions are among the most useful from which to derive chemical information and to provide
119 informative spectral fingerprints of the samples studied. In the case of food samples, for which
120 colour is a basic quality characteristic to be assessed, also the visible region is often considerably
121 informative.

122 HSI methods usually provide a considerable amount of data, structured in three dimensions (two
123 spatial and one spectral) and are often highly complex, especially when working in the near infrared
124 region, which is characterised by spectral overtone bands that often overlap considerably. In
125 addition, the effective application of HSI techniques in routine food controls is currently limited by
126 a series of problems, such as the presence of unwanted variations in the signals, due to factors
127 beyond our control.

128 In order to minimise these problems and to extract information relevant for analytical purposes, it is
129 essential to apply suitable signal processing and multivariate pattern recognition techniques [14].

130 In the last decades, great attention has been paid to the application of FT-IR spectroscopy in the
131 field of microbial safety, thanks both to the ability of the technique to resolve complex mixtures in
132 terms of composition and to the technological progress of modern instruments [15,16]. In particular,
133 the coupling of microscopic devices and MIR spectroscopy offered significant advantages for the
134 determination of compounds present at low concentration within heterogeneous matrices. Potential
135 of FTIR techniques for identification of contaminant bacteria in food has been systematically
136 investigated since the nineties of the 20th century, coupled with chemometric techniques, and
137 proved to be able to perform not only species but also strain differentiations [17,18]. Recent studies
138 efficiently coupled FTIR spectroscopy and microscopy for the characterisation of bacteria in fruit
139 juices, dairy products and meat [19,20]. It is worth remarking that most of the approaches described
140 in the literature include a separation step, aimed at extracting the bacterial cells from the complex
141 food matrices before analysis. This compromises the advantages of low invasiveness towards
142 sample and the performances of the method in terms of analysis time and cost.

143 In the present study, the capability of spectral and hyperspectral techniques to detect bacterial
144 surface contamination was investigated preliminarily on gel cultures, and subsequently on sliced
145 cooked ham. In more detail, two species of LAB were considered, namely *Lactobacillus curvatus*
146 and *Lactobacillus sakei*, both of which are responsible for common alterations in sliced cooked
147 ham.

148 LAB have been successfully studied by FTIR methods in the NIR and in the MIR spectral regions.

149 In particular, FTIR spectroscopy in the MIR range allowed the discrimination of the pure bacteria
150 analysed in transmission mode [21]. In the present study, the potential of three techniques were
151 investigated, with different equipment, respectively: a macroscopic hyperspectral scanner operating
152 in the NIR (10,470-5880 cm^{-1}) region, a FT-NIR spectrophotometer equipped with transmission
153 arm as sampling tool, working in the 12,500-5800 cm^{-1} region, and a FT-MIR microscopy operating
154 in the 4000-675 cm^{-1} region.

155 Multivariate exploratory data analysis was applied in order to extract useful information from the
156 data in order to verify the capability to distinguish the bacterial strains both in the culture and in the
157 food matrix.

158

159 **2. Materials and methods**

160 **2.1. Bacterial strains and culture conditions**

161 The study was carried out using *L. curvatus* VZ22 and *L. sakei* VZ35 strains, isolated from Varzi
162 PDO salami, and belonging to the bacterial collection of the Institute of Sciences of Food
163 Production of the National Research Council of Italy (CNR-ISPA). The strains were identified by
164 partial 16S rRNA sequencing according to McCabe and co-workers [22].

165 Before each experiment, the cultures were incubated overnight at 37 °C, in MRS broth (Scharlau
166 Microbiology, Barcelona, Spain). *Lactobacillus* strains were streaked out in triplicate on MRS agar
167 (Scharlau Microbiology), incubated at 37 °C for 72 h under anaerobic conditions (Anaerocult A,
168 Merck Millipore, Darmstadt, Germany) and submitted to spectral and hyperspectral analysis.

169 Visual appearance of Petri dishes is shown in Figure S1 (Supplementary Material), with agar gel
170 only (Figure S1.a) and with LAB cultures (Figures S1.b-c).

171 A commercial sliced cooked ham (about 2 mm thickness) was used for studying superficial
172 contamination by the same LAB species and it was purchased from an Italian supermarket. The
173 total load of bacteria of the slices before inoculum was $< 10^3$ cfu/g. A weighted amount (10 g) of
174 sample was serially diluted in one-quarter-strength Ringer's solution and plated in Aerobic Count
175 Plates (AC) Petrifilm (3M Canada, London, O Than, 1 mL of overnight cultures of *Lactobacillus*
176 strains (about 10^8 cfu/mL) was centrifuged and washed twice with 1 mL of phosphate buffered
177 saline (PBS), and then re-suspended in 1 mL of Ringer solution. Different spots within ham surface
178 were artificially contaminated by deposition of 100 μL LAB suspensions (about 10^7 cfu/spot). After
179 contamination, the solution was allowed to dry at room temperature and samples were immediately
180 submitted to spectroscopic analyses.

181

182 **2.2 Analytical techniques**

183 *2.2.1. FT-NIR spectroscopy*

184 FT-NIR measurements were performed using a Bruker Optic MPA FT-NIR spectrophotometer
185 equipped with transmission arm as sampling tool, working in the 12,500-5800 cm⁻¹ region. The
186 spectra were acquired in transmittance mode at 4 cm⁻¹ resolution as the average of 32 scans. A
187 rotating sample holder was used during spectra acquisition. This sampling procedure was necessary
188 since the FT-NIR spectrophotometer used was not equipped with sampling tools suitable for
189 punctual analysis. On the one hand, the rotation allows to eliminate the variation due to the intrinsic
190 heterogeneity of the sample, on the other hand, the collected signal results in a sort of average of the
191 characteristic signals of both LAB and agar gel. In fact, MRS agar was used as the chemical
192 substrate on which bacteria were inoculated. In order to account for potential spectral interferences
193 from the substrate, FT-NIR spectra were recorded both on the empty sample holder, as it is usually
194 done in spectroscopy, and on the Petri dish containing the MRS agar substrate. In more detail, two
195 series of spectra have been acquired on Petri dishes, considering two different backgrounds: an
196 empty glass Petri dish (E series) and a polymeric (polypropylene) Petri dish containing the gel
197 starter culture (G series). For each sample, two replicate measurements have been acquired for both
198 the E and G series in two measurement sessions, following a different random sequence. The two
199 different measurement sessions have been performed in the same day, by a time interval of less than
200 one hour to evaluate measurement repeatability. As for the artificially contaminated cooked ham,
201 the spectra were acquired under the same conditions used for Petri dishes, but the rotating sample
202 holder was removed. In fact, in this case, the rotation of the sample would be inappropriate, since
203 the contaminated area is located within a limited region of sample surface. An empty glass Petri
204 dish was used for background measurement. Two replicate spectra were acquired for each sample in
205 two measurement sessions, following a different random sequence.

207 *2.2.2. NIR hyperspectral imaging*

208 The hyperspectral imaging system consists of a desktop NIR Spectral Scanner (DV Optic)
209 embedding a Specim N17E reflectance imaging spectrometer, coupled to a Xenics XEVA 2608
210 camera (320 x 256 pixels). The system covers the spectral range from 900 to 1700 nm
211 (corresponding to 11110-5880 cm⁻¹) with a 5 nm resolution, for a total of 161 wavelengths.
212 Each image scene provided a data hypercube composed by 320 pixels per row and with a number of
213 rows of 283 with a spatial resolution of 0.47 mm. A dark silicon carbide (SiC) sandpaper sheet was
214 used as background for all the acquired images, since it is characterised by a very low and constant
215 reflectance spectrum. In the image scene also different reference materials, i.e., a high-reflectance

white ceramic tile and two ceramic tiles with different grey-scale tones, have been included for the possible image correction, as already done in [23].

For each Petri dish, three replicate images have been collected in random order, in addition to four background measurements taken at different times during the acquisition session in order to better evaluate the stability of the system over time.

The same artificially contaminated cooked ham were imaged by simply placing the slice onto a sandpaper sheet. Two replicate images were acquired for each sample in two measurement sessions, following a different random sequence.

2.2.3. FT-MIR micro spectroscopy

A Thermo Nicolet (Thermo Fisher Scientific, Waltham, MA, USA), iNTM10MX microscope, equipped with a mercury–cadmium–telluride detector cooled by liquid nitrogen, was used for spectroscopic analysis in the mid infrared region. The measurements were performed within the 4000-675 cm⁻¹ region, at a spectral resolution of 4 cm⁻¹, spanning an area of 200x200 μm², by 128 scans.

In the case for LAB cultures on Petri dishes, reflection absorption spectroscopy (RAS) measurements on metallic substrate (gilded glass slides) were performed, at a 45° incidence angle [24]. Three micro samples were taken from each Petri dish and applied onto a gilded substrate, performing three replicate measurements were performed on each sample, thus recording a total number of 54 spectra.

In the case for artificially contaminated cooked ham, an attenuated total reflection (ATR) imaging system equipped with a conical germanium crystal and purchased by Thermo Fisher Scientific (Waltham, MA, USA) was applied for the characterisation of sample surfaces. In particular, three different slices of ham were contaminated with both of the two LAB. Measurements were performed after evaporation of the bacteria suspensions deposited. Several spectra were recorded on the contaminated and non-contaminated areas for each ham slice with an objective aperture of 400x400 μm² corresponding to an investigation area of about 100x100 μm².

A dedicated software, OMNIC PictaTM (Thermo Fisher Scientific, Waltham, MA, USA), was applied for a preliminary manipulation of spectral data.

2.3. Data processing and analysis

2.3.1. Exploratory analysis

Principal component analysis (PCA) was applied for exploratory purposes to the FT-NIR and to the FT-MIR datasets, separately. In particular, both of the FT-NIR data matrices – obtained by

considering the two different backgrounds – have size $\{16 \times 1737\}$, while the FT-MIR data matrices have size $\{54 \times 1725\}$ (bacterial cultures on gel) and $\{746 \times 1725\}$ (bacteria on food matrix), respectively, where rows correspond to samples and columns correspond to the spectral variables recorded.

Several spectral pre-treatments, including Savitzki-Golay derivatives [25] and the standard normal variate (SNV) transform [26] were applied, both individually and in combination, in order to minimise unwanted signal variations, not related with chemical characterisation of bacterial samples.

As for the FT-NIR data, PCA was applied after column mean-centering, with the double aim of visualizing the data structure and detecting the presence of possible outliers (on the basis of the 99.7% confidence limit in the Q - T^2 plot).

In more detail, the detection of possible outliers is important since their presence can significantly affect the results of subsequent data analyses. In order to evaluate whether the samples identified as anomalous should be eliminated or are, instead, important to account for particular aspects of the system under study, two statistical parameters were considered: the distance of each sample from its projection on the hyperplane of the model, referred to as Q residual, and the distance of the projection of each sample on the PCA model from the centre of the model itself, referred to as Hotelling T^2 .

Let \mathbf{X} be a data matrix with size $\{n \times m\}$, obtained by measuring m variables on n samples. By means of PCA, the original data matrix \mathbf{X} is decomposed into three matrices, the first one accounting for the variance associated to samples (score matrix, \mathbf{T}), the second one accounting for the variance associated to variables (loading matrix, \mathbf{P}) and the third one describing the non-systematic variation (residuals matrix, \mathbf{E}) according to the following equation:

$$\mathbf{X} = \mathbf{TP}' + \mathbf{E} = \sum_{a=1}^A \mathbf{t}_a \mathbf{p}_a' + \mathbf{E} \quad (1)$$

Q residuals are calculated as the sum of squares of each row of the residuals matrix \mathbf{E} , according to the following equation for the i^{th} sample in \mathbf{X} , \mathbf{x}_i :

$$Q_i = \mathbf{e}_i \mathbf{e}_i' = \mathbf{x}_i (\mathbf{I} - \mathbf{PP}') \mathbf{x}_i' \quad (2)$$

where \mathbf{e}_i is the i^{th} row of \mathbf{E} , \mathbf{P} is the matrix of the A loading vectors considered in the PCA model and \mathbf{I} is the identity matrix of appropriate size.

T^2 corresponds to the sum of normalised squared score values:

$$T_i^2 = \sum_{a=1}^A \frac{t_{ia}^2}{s_a^2} = \frac{t_{i1}^2}{s_1^2} + \frac{t_{i2}^2}{s_2^2} + \dots + \frac{t_{iA}^2}{s_A^2} \quad (3)$$

281 where t_{ia} is the score value of the i th sample for the a^{th} PC and s_a^2 is the variance of the a^{th} score
282 vector, \mathbf{t}_a , in the model.

283 By describing the unusual variation of each sample outside the model, high Q values indicate
284 samples that do not follow the general pattern described by the PCA model. Conversely, the
285 Hotelling T^2 parameter quantifies the samples variability within the model and the contribution of
286 each sample in the definition of the model itself, so high T^2 values correspond to those samples
287 which have mainly influenced the model.

288

289 *2.3.2. Correction and segmentation of hyperspectral images*

290 Intensity values of the raw images were converted into the corresponding reflectance values by
291 applying a simple external calibration based on the high-reflectance standard reference and on the
292 dark current measured by covering the camera lens with its cap [27]. Due to the low S/N ratio of
293 spectra at the extremes of the measurement range, only the 150 spectral variables between 955 and
294 1700 nm (corresponding to 10470-5880 cm^{-1}) were considered for further analysis.

295 The first stage of the image processing (a preliminary exploratory data analysis step made by PCA
296 on several merged images) evidenced that the acquisition system is very stable over time, therefore
297 the images correction based on the reference materials included in the image scene was not
298 necessary. In view of these results and in order to reduce the computational load, all the
299 hyperspectral images were cropped, excluding the reference materials, to dimensions of 241×247
300 pixels.

301 PCA was then applied on each hyperspectral image, to segment the sample with respect to
302 background pixels. To this aim, the effectiveness of several data pretreatments such as the standard
303 normal variate (SNV) transform, detrending, first and second derivatives, column mean-centering
304 and column autoscaling, were evaluated both separately and in different combinations.

305

306 *2.3.3. Conversion of hyperspectral images into hyperspectrograms*

307 The two datasets of segmented images, one for the Petri dishes and one for the ham samples, were
308 converted into sets of one-dimensional signals, called hyperspectrograms, by means of a recently
309 developed algorithm [28], which is derived from the colourgram approach, previously developed
310 for the processing of RGB images [29–31].

311 This chemometric approach is aimed to significantly reduce the dataset size by compressing the
312 useful information contained within each image into an artificial signal created by merging in
313 sequence the frequency distribution curves of scores, Q residuals, Hotelling T^2 and loading vectors
314 obtained by applying PCA on the single unfolded hypercube data. Hyperspectrograms can then be

315 used as a compact set of descriptors and submitted to further multivariate analysis techniques. This
316 strategy also presents the great advantage of allowing for the simultaneous processing of tens up to
317 hundreds of hyperspectral images.

318 The choice of the appropriate number of PCs to be retained in the PCA models used for the
319 calculation of the hyperspectrograms does not represent a crucial point, since hyperspectrograms
320 may be undergo a variable selection step, in which the PCs accounting for non-informative
321 variability sources are discharged. In this case, a preliminary evaluation by PCA on a restricted
322 number of representative images is performed, to provide an estimate of the number of PCs
323 bringing useful information.

324 In Figure 1, the three average hyperspectrograms (corresponding to agar gel, *L. curvatus* and *L.*
325 *sakei* classes) obtained from the Petri dish images are reported together with the description of the
326 source of the various parts of the signals. In this case, a 5-PC model after column mean-centering as
327 the pre-treatment was considered.

328 PCA was applied on the whole datasets of hyperspectrograms to detect the presence of possible
329 outliers (the signals outside the 99.7% confidence limit in the $Q-T^2$ plot) and to visualise the data
330 structure, as for common NIR spectra.

331

332 **3. Results and discussion**

333 *3.1. FT-NIR spectra*

334 Visual investigation of FT-NIR spectra acquired for the E and G series evidenced a slight but
335 visible differentiation of the three classes of signals, in particular in the spectral region above 7500
336 cm^{-1} . The different series of spectra are reported in Figure S2 (Supplementary Material). It can be
337 noticed that signals marked as the E series (empty sample holder as background) are more intense
338 than those of the G series. For both of the series, signals corresponding to the three types of samples
339 analysed are visually different, in particular at wavenumbers higher than 7500 cm^{-1} .

340 For both of the series of spectra, the PCA model obtained from the mean-centered data did not
341 evidence the presence of outliers.

342 As for the E series, the PCA model (3 PCs, 97.7% explained variance) provided the PC1-PC2-PC3
343 score plot reported in Figure 2.a. In the 3-PCs space, the spectra corresponding to the *L. sakei* class
344 form a cluster clearly distinguishable (in particular, along PC2) with respect to the others; the *L.*
345 *curvatus* and the gel spectra are also distinguishable, though they appear quite close to each other;
346 the third principal component evidently accounts for the different analytical sessions, which were
347 separated by a time interval of less than one hour.

348 As for the G series, a 2-PC model was considered (96.6% explained variance). In Figure 2.b, the
349 PC1-PC2 score plot is reported; the samples belonging to the three classes appear grouped into
350 three distinct clusters. The cluster of *L. curvatus* appears more scattered in the PC space with
351 respect to the *L. sakei* one. As explained in paragraph 2.2.1, FT-NIR spectra were acquired by
352 averaging the signals collected during sample rotation. Looking at the corresponding RGB pictures
353 in Figure S1 (Supplementary Material), it can be noticed that *L. curvatus* colonies seem to be less
354 grown. This fact might be responsible for score distribution observed in the scatter plot: since *L.*
355 *curvatus* colonies tended to occupy smaller areas on the agar surface, they may have contributed
356 less to the average spectrum of each sample, which was also less repeatable. Comparing the models
357 obtained for the two series of FT-NIR spectra, a 3D-score plot was helpful to efficiently distinguish
358 signals belonging to the different types of samples for the E series, while a 2D-score plot was
359 sufficient for the G series. Moreover, the percentages of variance explained by the three PCs in the
360 case of the E series and by the two PCs in the case of the G series are very similar. These
361 considerations indicate that use of MRS agar background over the empty sample holder for signal
362 acquisition in future studies should be profitable.

363 Multivariate analysis of FT-NIR spectra recorded on cooked ham samples artificially contaminated
364 with LAB did not allow to efficiently detect the presence of bacteria. PCA was applied using
365 different spectra pre-treatments but, in each case, the three clusters corresponding to pure and
366 contaminated ham (with *L. curvatus* and *L. sakei*, respectively) in the score plots were not
367 distinguishable (*data not shown*).

368 This outcome could be due to different reasons, mainly concerning sensitivity of FT-NIR
369 measurements, which is lower with respect to other spectroscopic techniques such as FT-MIR [32].
370 About this, Cámara-Martos and co-workers analysed in transmission different LAB water solutions
371 in the 10^3 - 10^9 cfu/mL concentration range (the cell optical path used for the analysis of 1 mL of
372 solution was not specified) [33]. They were able to distinguish the different bacteria in solutions
373 containing high concentrations, but PCA was less effective for low concentrations.

374 In the present study, when 100 μ L of LAB suspension having concentration of 10^8 cfu/mL were
375 deposited over a flat surface, they were spread in a layer that is probably too thin (and so the LAB are
376 too diluted) to be revealed. In addition, in the present work, the spectrum of samples was composed
377 by the overlapping absorption bands of both ham and LAB but – at this level of dilution for the
378 contaminants – the contribution of ham may be overwhelming.

379 In the literature, a number of works about the assessment of microbial contamination of flesh by
380 using NIR-based methods, in particular hyperspectral imaging, can be found. In almost all of these
381 works, the amount of bacteria is estimated in meat and fish samples stored under controlled

conditions for different time intervals, ranging from a few hours to several days: under these conditions, samples undergo bacterial spoilage. Such a spoilage, which is responsible for off-flavours and slime formation, becomes apparent when microorganisms have reached concentrations around 10^7 - 10^8 CFU/g, depending on muscle type and pH [34]. At this point, food is definitely decayed from a sensory point of view and should be considered unsafe for human consumption. Concerning the use of classical FT-NIR spectroscopy, Lin and co-workers [35] applied short-wavelength near-infrared diffuse reflectance spectroscopy in the 600–1100 nm region to quantify microbial loads in chicken meat. Although they achieved satisfactory results, they remarked that sample spoilage produces biochemical changes in chicken meat (mainly proteolysis, indicated by the release of free amino acids), so that NIR measurements revealed the extent of biochemical changes in chicken – correlated to microbial spoilage – rather than directly measuring the presence of bacteria on breast muscle surface.

3.2. NIR hyperspectral images

Among the combinations of pre-treatments evaluated, detrending combined with column mean-centering provided the best distinction of the pixel cluster corresponding to the Petri dish from the clusters formed by background pixels. The score images obtained for all the sample images evidenced that the second PC allowed a clear distinction between the Petri dish and the background. In more detail, a threshold value of PC2 score equal to 0 was found to be useful to remove the background and the Petri dish border. As the final step, a further control aimed at manually removing extraneous pixels eventually not excluded was performed. As an example, the image of a *L. curvatus* sample, before and after segmentation, is reported in Figure 3.

A PCA model obtained on the hyperspectrogram dataset considering a number of significant PCs equal to 3 (accounting for about 65% of the total variance) did not reveal outliers. Looking at the PC1-PC2 score plot (Figure 4), three clusters corresponding to the three different classes were distinguished; in particular, PC1 allowed a clear separation of gel from bacteria samples, while PC2 allowed to distinguish between the two different bacterial species. Since the hyperspectrogram approach allows to condense the information of a whole image into a single signal, it can be used to estimate the actual repeatability of replicate images. In Figure 4, the replicates of the same samples were indicated by using the same symbol, and this allowed to verify, with few exceptions, the reproducibility of the system. This finding suggests that performing PCA on hyperspectrograms could be also considered a valuable alternative method to evaluate the instrumental stability over time, for instance using standard images, or to investigate a dynamic process such as sample ageing.

415 Also in this case, analysis of hyperspectrograms obtained from the segmented images acquired on
416 artificially contaminated ham did not evidenced neither the presence nor the type of bacteria (*data*
417 *not shown*). The possible reasons for this failure are similar to those invoked in the case of FT-NIR
418 spectra: absorption bands of ham may be overwhelming with respect to the LAB ones, and the
419 concentration of bacteria over the ham surface is too low to be revealed. In addition,
420 instrumentation for the acquisition of hyperspectral images does not explore the entire NIR spectral
421 range.

422 Very recent research works proved that imaging techniques would allow to estimate the extent of
423 bacterial contamination on meat and fish samples. As an example, He et al. [36] assessed LAB
424 values in salmon fillets during a 12-day cold spoilage process. Their study confirmed that no typical
425 absorption band of LAB was found in the spectral profiles, but spectral data of salmon samples
426 could be mined – by using suitable chemometric tools – to indirectly predict LAB, as the growth of
427 LAB is closely related with chemical composition of salmon flesh. In the work by Tao et al. [37], a
428 particular hyperspectral imaging system was used, namely the hyperspectral scattering system, in
429 which point light is applied as the illuminant source, in a manner that the backscattering image of
430 samples is captured. During the meat spoilage process, meat microstructure undergoes changes,
431 together with bacterial load and chemical composition. These changes could be reflected by the
432 derived light absorption and scattering features from the hyperspectral scattering image. Barbin and
433 co-workers [38] investigated by HSI the microbial growth in pork meat samples stored under
434 refrigerated conditions during 21 days. They were able to classify samples as fresh or spoiled, and
435 the comparison of the spectral features of the two sample categories allowed to determine how
436 microbial spoilage affects spectral fingerprints of pork meat. The main differences were observed in
437 the wavelength range between 1300 nm and 1600 nm, where N–H stretching of proteins (amines
438 and amides) and their interactions with water are observed: this suggests the occurrence of
439 proteolytic changes, which are recognised as the main indicator for the starting of spoilage in meat
440 products.

441 Concerning the study of bacterial surface contamination of sliced cooked ham, the focus of the
442 present study was on early detection of LAB presence before ham undergoes evident chemical and
443 physical changes: for this reason, samples were analysed just after evaporation of LAB suspensions
444 deposited. Therefore, the failure of NIR-based methods to evidence the presence of LAB on ham
445 surface are potentially due to the lack of spoilage of the matrix, which is – conversely – a key aspect
446 in all of the studies mentioned above.

447

3.3. FT-MIR micro spectroscopy

FTIR microscopy in the MIR region presents the disadvantage of requiring an appropriate sampling both of the bacterial cultures and of the food matrix contaminated with bacteria. However, working at the microscopic scale, it is possible to obtain spectra related to small area thus allowing to maximise the signal of bacteria, with respect to that of the matrix, in the points in which they are present. Accordingly to what was observed by Oust and co-workers [21], Figure S2 (Supplementary Material) shows the average profile of FT-MIR RAS spectra recorded for *L. curvatus* (FigureS2.a) and *L. sakei* (Figure S2.b) from gel cultures, respectively. The main bands are ascribable to amide N-H stretching (3200 cm^{-1}), amide I and II bands (near 1650 cm^{-1} and 1540 cm^{-1} , respectively), related to the protein pattern [39], and to C=O stretching of lipid esters (shoulder visible at 1743 cm^{-1} in Figure S2.b) [19]. As a matter of facts, the whole protein pattern has been reported as a key difference between the two species [40,41].

PCA applied to the whole dataset (27 spectra recorded for each bacterial species) shows that the MIR spectra recorded contain information useful to identify the two *Lactobacillus* species.

In particular, Figure 5.a reports the PC1-PC3 score plot obtained after application of SNV transform, first derivative and column mean-centering. Two well defined clusters, corresponding to *L. sakei* (blue) and *L. curvatus* (red), respectively, are evident, even though an important dispersion among analytical replicates is noticeable. Figure 5.b show the corresponding loading plot, in which it is possible to identify spectral variables around $1620\text{--}1650\text{ cm}^{-1}$ (amide I band) and 1740 cm^{-1} (ester band) as the most involved in the characterisation of the two bacterial species.

Analysis of FT-MIR reflection spectra recorded on cooked ham samples artificially contaminated with LAB showed that signals in the mid-infrared region are useful to detect and identify bacterial species also on the surface of real food matrices.

Outcomes of PCA, after application of SNV transform, second derivative and column mean-centering, are shown in Figure 6. The PC1-PC2 score plot (Figure 6.a) clearly shows a differentiation between non-contaminated areas (blue scores) and portions of ham surface contaminated with *L. curvatus* (red scores) and *L. sakei* (green scores), respectively.

From a joint analysis of the corresponding loading plot (Figure 6.b), the main role of ester and amide bands (at around 1740 and 1620 cm^{-1} , respectively) for the characterisation of bacteria is confirmed. Simultaneously, a band around 2916 cm^{-1} , ascribable to C-H stretching emerges as a marker for the differentiation of the two LAB species on ham

4. Conclusions

481 The results obtained in this study demonstrate that the spectroscopic techniques investigated may
482 represent an effective screening tool to detect surface bacterial contamination in samples and, in
483 particular, to recognise the species to which bacteria belong. All the techniques applied proved to be
484 effective when LAB are laid down on a simple matrix, such as MRS agar. Concerning the Petri dish
485 samples, both FT-NIR and FT-MIR techniques gave very satisfactory results on spectra treated
486 using common signal pre-processing methods. In the case of hyperspectral images, the datacubes
487 were converted into hyperspectrograms before performing exploratory analysis by PCA.
488 Conversion of images into hyperspectrograms is appropriate, as demonstrated by the encouraging
489 results obtained. In particular, this novel approach has proved to be effective to extract the useful
490 information from the images and, at the same time, to disregard the uninformative one.

491 When LAB are inoculated on a more complex matrix, such as sliced cooked ham, not all of the
492 spectroscopic techniques worked equally well. Preliminary tests indicated that FT-MIR microscopy
493 may provide satisfactory responses when dealing with real food samples contaminated at spoilage
494 concentration, also thanks to the fact that it allows localised investigations to be performed. On the
495 contrary, neither FT-NIR nor NIR hyperspectral imaging were effective in detecting the presence
496 and the type of bacteria. The main reason for these outcomes may be ascribable to the limited
497 sensitivity of NIR-based spectroscopic techniques with respect to FT-MIR.

498 These preliminary results suggest that a more in-depth investigation is required, which could be
499 interesting because NIR-based techniques have many advantages in terms of practical applications
500 and implementation of routine control systems. In particular, FT-NIR measurements are very fast –
501 though not so sensitive – so they could be used for screening purposes; when the spectrophotometer
502 is equipped with fiber optic probes, also online screening is possible. Further advantages connected
503 to NIR imaging implementations involve the possibility to visualise the location of contaminated
504 areas on the sample. The hyperspectrogram approach, in which each image is converted into a
505 signal, still allows the spatial localisation of the contamination through the reconstruction of false
506 color images after multivariate data analysis [42]. In addition, the conversion of each image into a
507 signal of few hundreds of points enables the simultaneous modelling of tens (but it also works for
508 hundreds) of hyperspectral images. In this way, NIR imaging might be implemented as a suitable
509 system for monitoring different species of bacterial contaminants directly in industrial plants.

510 FTIR microscopy in the MIR region seems a promising tool for the identification of contaminated
511 food matrix. However, being a micro chemical technique just a small area at a time can be analysed,
512 and a proper sampling strategy is needed. Thus, this approach may find application to confirm or
513 refine preliminary screening data obtained with macro techniques such as FT-NIR and NIR
514 hyperspectral analyses.

515

516 **References**

- 517 [1] CE Regulation n. 1333/2008.
- 518 [2] DL 109/1992.
- 519 [3] CE Regulation n. 1169/2011.
- 520 [4] CE Regulation n. 1441/2007.
- 521 [5] P. Hu, G. Zhou, X. Xu, C. Li, Y. Han, Characterization of the predominant spoilage bacteria
522 in sliced vacuum-packed cooked ham based on 16S rDNA-DGGE, *Food Control*. 20 (2009)
523 99–104. doi:10.1016/j.foodcont.2008.02.007.
- 524 [6] C. Cantoni, S. Milesi, L. Iacumin, Bacterial alterations of cooked packaged hams and the
525 responsibility of *Enterococcus faecalis* in these defect, *Ind. Aliment.* 48 (2009) 31–35.
- 526 [7] A. Gowen, C. O'Donnell, P. Cullen, G. Downey, J. Frias, Hyperspectral imaging – an
527 emerging process analytical tool for food quality and safety control, *Trends Food Sci.*
528 *Technol.* 18 (2007) 590–598. doi:10.1016/j.tifs.2007.06.001.
- 529 [8] H. Cen, Y. He, Theory and application of near infrared reflectance spectroscopy in
530 determination of food quality, *Trends Food Sci. Technol.* 18 (2007) 72–83.
531 doi:10.1016/j.tifs.2006.09.003.
- 532 [9] I. Kavdir, R. Lu, D. Ariana, M. Ngouajio, Visible and near-infrared spectroscopy for
533 nondestructive quality assessment of pickling cucumbers, *Postharvest Biol. Technol.* 44
534 (2007) 165–174. doi:10.1016/j.postharvbio.2006.09.002.
- 535 [10] D.W. Sun, *Hyperspectral Imaging for Food Quality Analysis and Control*, Elsevier Inc.,
536 2010.
- 537 [11] J.M. Amigo, H. Babamoradi, S. Elcoroaristizabal, Hyperspectral image analysis. A tutorial,
538 *Anal. Chim. Acta*. 896 (2015) 34–51. doi:10.1016/j.aca.2015.09.030.
- 539 [12] J.M. Amigo, Practical issues of hyperspectral imaging analysis of solid dosage forms., *Anal.*
540 *Bioanal. Chem.* 398 (2010) 93–109. doi:10.1007/s00216-010-3828-z.
- 541 [13] C. Esquerre, A. Gowen, G. Downey, C. O'Donnell, Wavelength selection for development
542 of a near infrared imaging system for early detection of bruise damage in mushrooms
543 (*Agaricus bisporus*), *J. Near Infrared Spectrosc.* 20 (2012) 537. doi:10.1255/jnirs.1014.
- 544 [14] J. Burger, A. Gowen, Data handling in hyperspectral image analysis, *Chemom. Intell. Lab.*
545 *Syst.* 108 (2011) 13–22. doi:10.1016/j.chemolab.2011.04.001.
- 546 [15] D. Alexandrakis, G. Downey, A.G.M. Scannell, Rapid Non-destructive Detection of
547 Spoilage of Intact Chicken Breast Muscle Using Near-infrared and Fourier Transform Mid-
548 infrared Spectroscopy and Multivariate Statistics, *Food Bioprocess Technol.* 5 (2009) 338–
549 347. doi:10.1007/s11947-009-0298-4.
- 550 [16] F. Valerio, P. De Bellis, M. Di Biase, S.L. Lonigro, B. Giussani, A. Visconti, et al.,
551 Diversity of spore-forming bacteria and identification of *Bacillus amyloliquefaciens* as a
552 species frequently associated with the ropy spoilage of bread., *Int. J. Food Microbiol.* 156
553 (2012) 278–85. doi:10.1016/j.ijfoodmicro.2012.04.005.
- 554 [17] D. Naumann, V. Fijala, H. Labischinski, The differentiation and identification of pathogenic
555 bacteria using FT-IR and multivariate statistical analysis, *Mikrochim. Acta*. 94 (1988) 373–
556 377. doi:10.1007/BF01205910.

- 557 [18] D. Naumann, D. Helm, H. Labischinski, Microbiological characterizations by FT-IR
558 spectroscopy., *Nature*. 351 (1991) 81–2. doi:10.1038/351081a0.
- 559 [19] R. Davis, J. Irudayaraj, B.L. Reuhs, L.J. Mauer, Detection of *E. coli* O157:H7 from ground
560 beef using Fourier transform infrared (FT-IR) spectroscopy and chemometrics., *J. Food Sci.*
561 75 (2010) M340–6. doi:10.1111/j.1750-3841.2010.01686.x.
- 562 [20] R. Davis, Y. Burgula, J. Irudayaraj, B.L. Reuhs, L.J. Mauer, Fourier Transform Infrared
563 (FT-IR) Spectroscopy Coupled with Filtration and Immunomagnetic Separation for the
564 Detection of *Escherichia coli* O157:H7 in Ground Beef, in: 2010 18th Bienn. Univ.
565 Micro/Nano Symp., IEEE, 2010: pp. 1–6. doi:10.1109/UGIM.2010.5508916.
- 566 [21] A. Oust, T. Mørtrø, C. Kirschner, J.A. Narvhus, A. Kohler, FT-IR spectroscopy for
567 identification of closely related lactobacilli., *J. Microbiol. Methods*. 59 (2004) 149–62.
568 doi:10.1016/j.mimet.2004.06.011.
- 569 [22] K.M. McCabe, G. Khan, Y.-H. Zhang, E.O. Mason, E.R.B. McCabe, Amplification of
570 Bacterial DNA Using Highly Conserved Sequences: Automated Analysis and Potential for
571 Molecular Triage of Sepsis, *Pediatrics*. 95 (1995) 165–169.
- 572 [23] A. Ulrici, S. Serranti, C. Ferrari, D. Cesare, G. Foca, G. Bonifazi, Efficient chemometric
573 strategies for PET–PLA discrimination in recycling plants using hyperspectral imaging,
574 *Chemom. Intell. Lab. Syst.* 122 (2013) 31–39. doi:10.1016/j.chemolab.2013.01.001.
- 575 [24] S. Prati, M. Quaranta, G. Sciutto, I. Bonacini, L. Litti, M. Meneghetti, et al., Use of nano
576 gold obtained by laser ablation for SEIRA analyses of colorants, *Herit. Sci.* 2 (2014) 28.
577 doi:10.1186/s40494-014-0028-8.
- 578 [25] A. Savitzky, M.J.E. Golay, Smoothing and Differentiation of Data by Simplified Least
579 Squares Procedures, *Anal. Chem.* 36 (1964) 1627–1639. doi:10.1021/ac60214a047.
- 580 [26] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Standard Normal Variate Transformation and De-
581 trending of Near-Infrared Diffuse Reflectance Spectra, *Appl. Spectrosc.* 43 (1989) 772–777.
582 doi:10.1366/0003702894202201.
- 583 [27] J. Burger, P. Geladi, Hyperspectral NIR image regression part I: calibration and correction,
584 *J. Chemom.* 19 (2005) 355–363. doi:10.1002/cem.938.
- 585 [28] C. Ferrari, G. Foca, A. Ulrici, Handling large datasets of hyperspectral images: reducing data
586 size without loss of useful information., *Anal. Chim. Acta.* 802 (2013) 29–39.
587 doi:10.1016/j.aca.2013.10.009.
- 588 [29] A. Antonelli, M. Cocchi, P. Fava, G. Foca, G.C. Franchini, D. Manzini, et al., Automated
589 evaluation of food colour by means of multivariate image analysis coupled to a wavelet-
590 based classification algorithm, *Anal. Chim. Acta.* 515 (2004) 3–13.
591 doi:10.1016/j.aca.2004.01.005.
- 592 [30] G. Foca, F. Masino, A. Antonelli, A. Ulrici, Prediction of compositional and sensory
593 characteristics using RGB digital images and multivariate calibration techniques., *Anal.*
594 *Chim. Acta.* 706 (2011) 238–45. doi:10.1016/j.aca.2011.08.046.
- 595 [31] A. Ulrici, G. Foca, M.C. Ielo, L.A. Volpelli, D. Pietro Lo Fiego, Automated identification
596 and visualization of food defects using RGB imaging: Application to the detection of red
597 skin defect of raw hams, *Innov. Food Sci. Emerg. Technol.* 16 (2012) 417–426.
598 doi:10.1016/j.ifset.2012.09.008.
- 599 [32] M. Khanmohammadi, S. Armenta, S. Garrigues, M. de la Guardia, Mid- and near-infrared
600 determination of metribuzin in agrochemicals, *Vib. Spectrosc.* 46 (2008) 82–88.

doi:10.1016/j.vibspec.2007.10.005.

[33] F. Cámara-Martos, G. Zurera-Cosano, R. Moreno-Rojas, R.M. García-Gimeno, F. Pérez-Rodríguez, Identification and Quantification of Lactic Acid Bacteria in a Water-Based Matrix with Near-Infrared Spectroscopy and Multivariate Regression Modeling, *Food Anal. Methods*. 5 (2011) 19–28. doi:10.1007/s12161-011-9221-5.

[34] F.J.G. Schreurs, Post-mortem changes in chicken muscle, *Worlds. Poult. Sci. J.* 56 (2000) 319–346. doi:10.1079/WPS20000023.

[35] M. Lin, M. Al-Holy, M. Mousavi-Hesary, H. Al-Qadiri, A.G. Cavinato, B.A. Rasco, Rapid and quantitative detection of the microbial spoilage in chicken meat by diffuse reflectance spectroscopy (600–1100 nm), *Lett. Appl. Microbiol.* 39 (2004) 148–55. doi:10.1111/j.1472-765X.2004.01546.x.

[36] H.-J. He, D.-W. Sun, D. Wu, Rapid and real-time prediction of lactic acid bacteria (LAB) in farmed salmon flesh using near-infrared (NIR) hyperspectral imaging combined with chemometric analysis, *Food Res. Int.* 62 (2014) 476–483. doi:10.1016/j.foodres.2014.03.064.

[37] F. Tao, Y. Peng, C.L. Gomes, K. Chao, J. Qin, A comparative study for improving prediction of total viable count in beef based on hyperspectral scattering characteristics, *J. Food Eng.* 162 (2015) 38–47. doi:10.1016/j.jfoodeng.2015.04.008.

[38] D.F. Barbin, G. ElMasry, D.-W. Sun, P. Allen, N. Morsy, Non-destructive assessment of microbial contamination in porcine meat using NIR hyperspectral imaging, *Innov. Food Sci. Emerg. Technol.* 17 (2013) 180–191. doi:10.1016/j.ifset.2012.11.001.

[39] A. Alvarez-Ordóñez, D.J.M. Mouwen, M. López, M. Prieto, Fourier transform infrared spectroscopy as a tool to characterize molecular composition and stress response in foodborne pathogenic bacteria., *J. Microbiol. Methods*. 84 (2011) 369–78. doi:10.1016/j.mimet.2011.01.009.

[40] J. Samelis, E. Tsakalidou, J. Metaxopoulos, G. Kalantzopoulos, Differentiation of *Lactobacillus sake* and *Lact. curvatus* isolated from naturally fermented Greek dry salami by SDS-PAGE of whole-cell proteins, *J. Appl. Bacteriol.* 78 (1995) 157–163. doi:10.1111/j.1365-2672.1995.tb02836.x.

[41] M.J. Benito, M.J. Serradilla, S. Ruiz-Moyano, A. Martín, F. Pérez-Nevado, M.G. Córdoba, Rapid differentiation of lactic acid bacteria from autochthonous fermentation of Iberian dry-fermented sausages., *Meat Sci.* 80 (2008) 656–61. doi:10.1016/j.meatsci.2008.03.002.

[42] C. Ferrari, G. Foca, R. Calvini, A. Ulrici, Fast exploration and classification of large hyperspectral image datasets for early bruise detection on apples, *Chemom. Intell. Lab. Syst.* 146 (2015) 108–119. doi:10.1016/j.chemolab.2015.05.016.

637 **Figure Captions**

638

639 **Figure 1.** Profiles of the average hyperspectrograms obtained for the three classes.

640

641 **Figure 2.** Scores plots obtained for the E (a) and the G (b) series. White symbols for gel samples, red for *L.*
642 *curvatus* and blue for *L. sakei*; triangles and circles distinguish the analysis sessions.

643

644 **Figure 3.** Average intensity gray-scale image of a *L. curvatus* sample before (a) and after (b) segmentation.

645

646 **Figure 4.** Scores plot obtained from PCA applied to the Petri dishes hyperspectrograms dataset. White
647 symbols identify gel samples, red *L. curvatus* and blue *L. sakei*. Different symbols indicate different
648 measurement replicates on the same samples.

649

650 **Figure 5.** PC1-PC3 score (a) and loading (b) plots of FT-MIR RAS spectra recorded for *L. curvatus* (red
651 scores) and *L. sakei* (blue scores) on agar gel.

652

653 **Figure 6.** PC1-PC2 score (a) and loading (b) plots of FT-MIR reflection spectra recorded for cooked ham
654 (green scores), and cooked ham artificially contaminated with *L. curvatus* (red scores) and *L. sakei* (blue
655 scores), respectively.

656

657

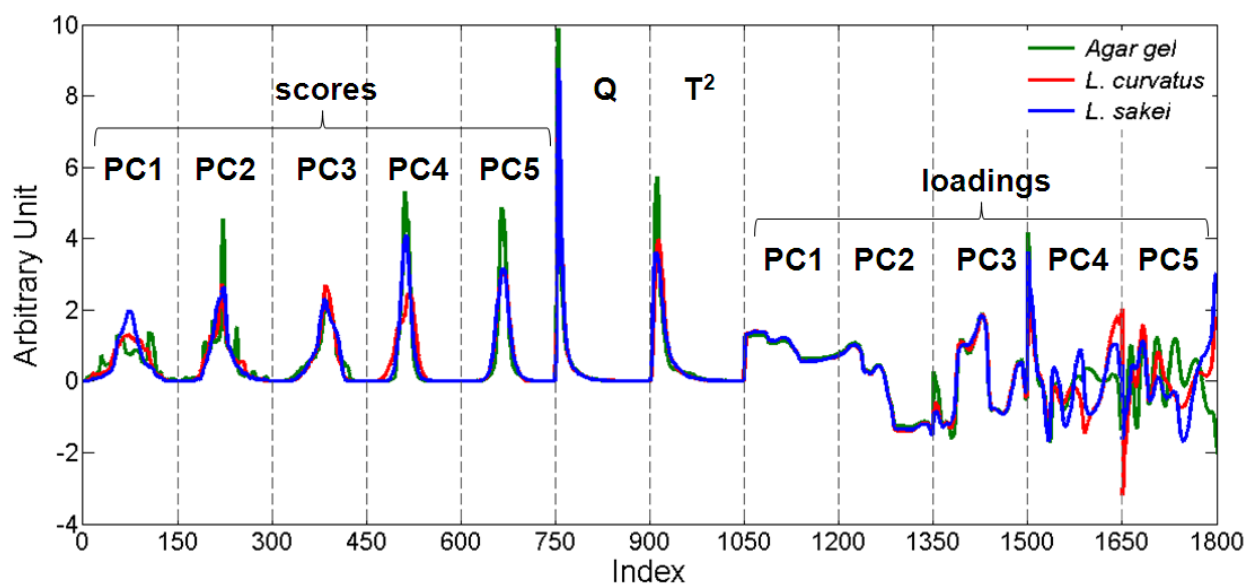


Figure 1

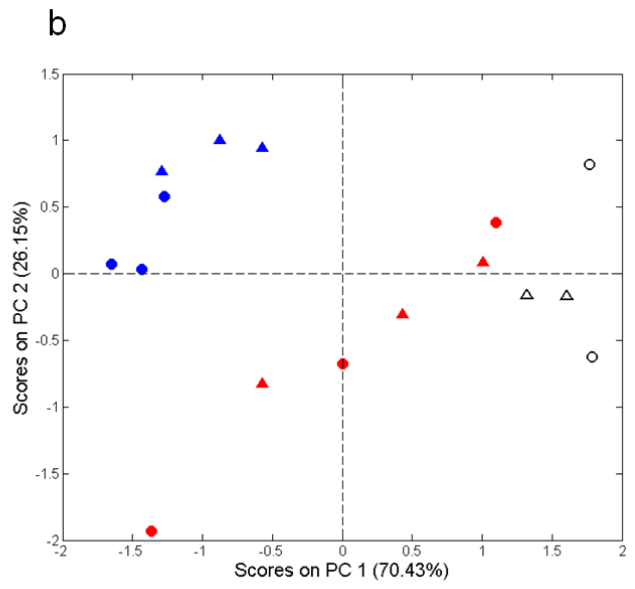
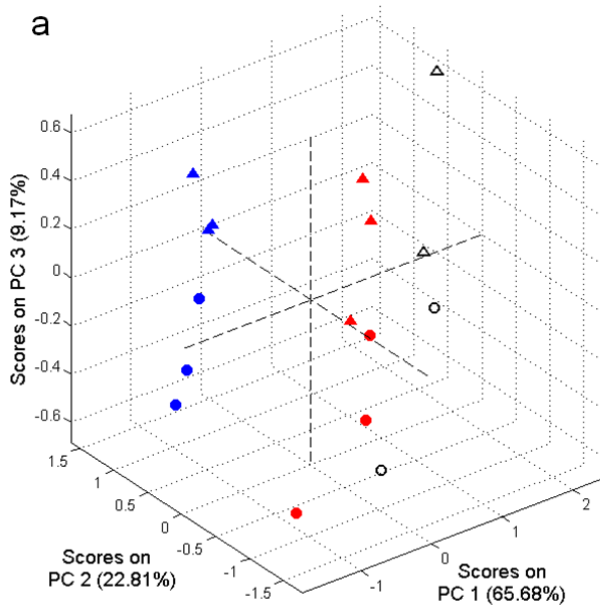


Figure 2

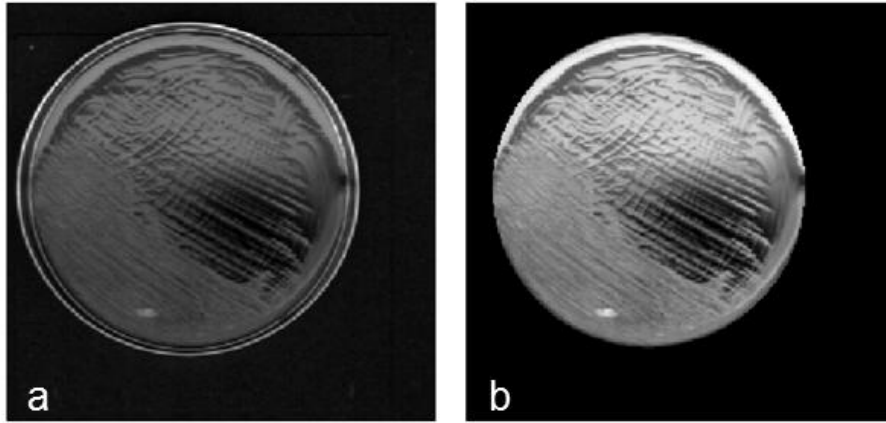


Figure 3

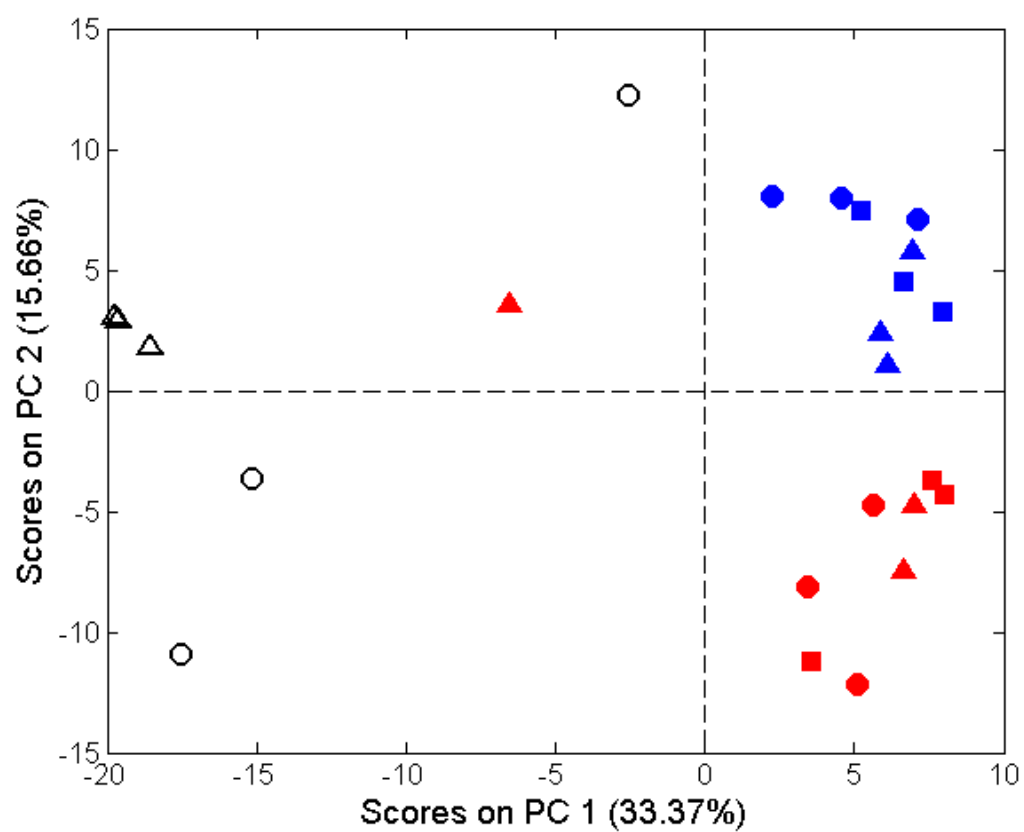


Figure 4

Figure 1 is a PCA plot showing the separation of three groups of *Arabidopsis thaliana* accessions based on 1000 SNPs. The x-axis represents the first principal component (PC1), accounting for 80.7% of the genetic variation. The y-axis represents the second principal component (PC2), accounting for 10.1% of the variation. The groups are color-coded: green for *cit2*, blue for *Vz35-1*, and red for *Vz22-3*. The *cit2* group is clustered in the upper left, *Vz35-1* in the middle right, and *Vz22-3* in the lower left. The plot is labeled 'a' in the top right corner.

b

This loading plot displays the distribution of metabolite variables along the first two principal components. The x-axis represents Loadings on PC1 (80.7%), ranging from -0.1 to 0.25. The y-axis represents Loadings on PC2 (10.1%), ranging from -0.1 to 0.15. Metabolites are labeled with their retention times (min). Variables like 1706.97, 1704.84, and 1710.62 show high positive PC2 loadings. A large cluster of variables is centered around the origin. Variables such as 1739.55, 1741.48, and 1743.41 show high positive PC1 loadings.

Figure 6