International Journal of Food Microbiology 238 (2016) 1–6

Contents lists available at ScienceDirect

International Journal of Food Microbiology

journal homepage: www.elsevier.com/locate/ijfoodmicro

Assessment of table olive fermentation by functional data analysis

M.A. Ruiz-Bellido [a], V. Romero-Gil [a,b], P. García-García [b], F. Rodríguez-Gómez [b], F.N. Arroyo-López [b,*], &, A. Garrido-Fernández [b]

[a] Regulatory Council of PDO Aloreña de Málaga table olives, C/Dehesa, 80, 29560 Pizarra, Málaga, Spain
[b] Food Biotechnology Department, Instituto de la Grasa (CSIC), University Campus Pablo de Olavide, Building 46. Ctra. Utrera, km 1, 41013 Seville, Spain

ARTICLE INFO

Article history:
Received 15 May 2015
Received in revised form 20 May 2016
Accepted 21 August 2016
Available online 23 August 2016

Keywords:
Aloreña de Málaga
Smoothing
Functional regression
Functional permutation tests
R software
Yeasts

ABSTRACT

For the first time, functional data analysis (FDA) was used to assess the effects of different treatments on Protection Denomination of Origin *Aloreña de Málaga* table olive fermentations, focusing on the evolution of yeast population. The analysis of fermentation by a conventional approach led to scarce information. However, the transformation of microbial (and also physicochemical) data into smooth curves allowed the application of a new battery of statistical tools for the analysis of fermentations (functional pointwise estimation of the averages and standard deviations, maximum, minimum, first and second derivatives, functional regression, and functional F and t-tests). FDA showed that all the treatments assayed led to similar trends in yeast population while changes in pH and titratable acidity profiles led to several significant differences. Therefore, FDA represents a promising and valuable tool for studying table olive fermentations and for food microbiology in general.

1. Introduction

Table olive fermentation, one of the most important fermented vegetables of the Mediterranean basin, implies changes in several physicochemical and microbiological parameters over time (Garrido-Fernández et al., 1997). Usually, these changes are monitored, and the data subjected to diverse statistical approaches, particularly the fitting of appropriate kinetic models (Arroyo-López et al., 2010). Then, the effects of the environmental variables on the estimated parameters are analysed by ANOVA, Response Surface Methodology, logistic regression, or multivariate methods (Arroyo-López et al., 2010; Bautista-Gallego et al., 2013a; Bevilacqua et al., 2009). Specifically, for the study of microbial populations during fermentation of *Aloreña de Málaga*, a table olive speciality with Protection Denomination of Origin (PDO) in Spain, several primary and secondary predictive models have been applied. This way, the microbial decay and growth data were modelled by Weibull survival and Gompertz models, respectively (Arroyo-López et al., 2007). A quasi-chemical primary model fitted lactic acid bacteria (LAB) and yeast populations during *Aloreña de Málaga* storage not only during growth but also throughout the declining phase (Echevarria et al., 2010). The physicochemical parameter changes occurring during olive fermentation can also follow diverse kinetic models. Pseudo-second and first order decay models fit the changes in sugars, polyphenols, and colour from cracked *Aloreña de Málaga* olives (Arroyo-López et al., 2007, 2008; Bautista-Gallego et al., 2011) while a third-order kinetic model fits the lactic acid production in a mixture of diverse chloride salts (Bautista-Gallego et al., 2010). However, microbiological (or physicochemical) data obtained from fermentation processes, and especially from table olives, cannot always be satisfactorily fitted by mathematical models, a circumstance that could make the quantitative study of the process and the comparison among treatments difficult.

Although functional data models have been a rather common technique in statistics, the term functional data analysis (FDA) was popularized by Ramsay and Dalzell (1991). FDA is related to the representation, summarization, and analysis of data obtained from curves. Thereby, FDA considers the data obtained from a series of samples over time as a curve or function. The concept is, then, clearly applicable to table olive fermentation because the microbiological (or physicochemical) data from the successive samples may represent the actual curve of their changes over time. The theoretical and practical aspects of FDA may be found in Ramsay and Silverman (1997, 2002, 2005) and Ramsay et al. (2009). Furthermore, Ramsay et al. (2014) also implemented an R package ("fda") for the analysis of data. Their studies are complementary to those carried out by Ferraty and Vieu (2006) on non-parametric data analysis. In the field of food technology, Bi and Kuesten (2013) have applied FDA to investigate the sensory intensity-time data, and their work represents an invaluable contribution for the diffusion of this methodology. Thereby, this statistical tool has very diverse applications. It has been used to study the degradation by the liver of chylomicron remnants, which excess may contribute to atherosclerosis (Nuzzo, 2002), or to assess the effects of different rearing environments on a population of asymptotic growth curves of Brown Kiwi (Jones et al., 2009). FDA of variance (fANOVA) was applied to study the effect of cultivar origin and shelf-life exposure time on the NIR apple spectra

* Corresponding author.
E-mail address: fnarroyo@cica.es (F.N. Arroyo-López).

http://dx.doi.org/10.1016/j.ijfoodmicro.2016.08.031
0168-1605/

(Bobelyn et al., 2010). Functional Principal Component Analysis (fPCA) applied to voltammetric data (and their first derivative) from urea and melamine-adulterated and non-adulterated milk samples allowed their correct classification by K-nearest neighbours (Hilding-Ohlsson et al., 2012). Ferraty et al. (2007) used a spectrometric data set (developed to control the fat content) for the factor-based comparison of groups while Aguilera et al. (2013) have reviewed the FDA of chemometric data (spectrum of meat or NIR spectrum) to predict the oil content in corn samples. However, FDA has not been yet applied to the study of the microbial populations during food fermentations. Therefore, the application of FDA to study table olive processing would represent an important step in extending its application to vegetable fermentation and also to food microbiology in general.

The aim of the present work was the application of FDA (vs. conventional approach) to investigate the changes in the yeast population, as well as to their associated physicochemical data, during the fermentation of PDO *Aloreña de Málaga* table olive subjected to diverse brining conditions.

2. Material and methods

2.1. Experimental design

The study was performed with *Aloreña de Málaga* fruits, harvested at the green ripe stage during the 2013/14 season (Valle del Guadalhorce, Málaga, Spain). For the experiments, 154 kg of whole fruits (cured type), or cracked olives (traditional type), were placed in 220 l containers (drums), where they were subjected to spontaneous fermentation after brining in 66 l of the following solutions: i) CC treatment (usual brine conditions of cured olives): 7 g/100 ml NaCl, 0.1 g/100 ml citric acid (CA), 0.5 g/100 ml acetic acid (AA); ii) CI treatment (highly acidified, cured olives): no salt, 0.1 g/100 ml CA, 1.6 g/100 ml AA; iii) CII treatment (moderately acidified, cured olives): no salt, 0.1 g/100 ml CA, 1.0 g/100 ml AA; iv) CT treatment (usual brine conditions of cracked, traditional olives): 11 g/100 ml NaCl solution, and v) RT treatment (fruits cracked after 72 h respiration at room temperature): brined in an 11 g/100 ml NaCl solution. All the treatments (5 fermentation systems) were run in duplicate. The containers were covered with their lids and stored at room temperature in the factory S.C.A. Copusan (Alozaina, Málaga, Spain). At different sampling times (1, 15, 38, 52, 80, 137, 250 and 380 days), 15 ml samples were aseptically withdrawn from the centre of both replicates of each treatment (drums) for their microbiological and physicochemical analysis. When necessary, the removed brine was replaced with the corresponding original brine. Then, the drums were covered again with the lid till the next sampling.

2.2. Microbiological analyses

Brine samples were diluted, if necessary, in a sterile saline solution (0.9 g/100 ml NaCl), and plated using a Spiral System model dwScientific (Dow Whitley Scientific Limited, England) on appropriate medium. *Enterobacteriaceae* were enumerated on VRBD (Crystal-violet Neutral-Red bile glucose)-agar (Merck, Darmstadt, Germany), LAB on MRS (de Man, Rogosa and Sharpe)-agar (Oxoid) with 0.02% sodium azide (Sigma, St. Louis, USA), and yeasts on YM (yeast-malt-peptone-glucose medium)-agar (Difco™, Becton and Dickinson Company, Sparks, MD, USA) supplemented with oxytetracycline and gentamicin sulphate as selective agents for yeasts. The plates were incubated at 30 °C for 24 (*Enterobacteriaceae)* or 48 (LAB and yeasts) hours and counted using a Flash & Go (IUL, Barcelona, Spain) image analysis system. Counts were expressed as $\log_{10}$ cfu/ml.

2.3. Physicochemical analyses

The analyses for pH and titratable acidity in the cover brines were carried out using the standard methods developed for table olives (Garrido-Fernández et al., 1997) using a Titroprocessor mod 670 (Metrohm Instrument, Herisau, Switzerland).

2.4. Statistical analysis

For the analysis of the microbial populations by a conventional approach, the model of the two-term Gompertz equation proposed by Bello and Sánchez Fuertes (1995) for microbial growth/decline was used to fit the data of the yeast population changes over time. Similarly, several kinetic models were also checked for fitting the pH and titratable acidity changes throughout fermentation, being the three parameter formation kinetic model ($y = a + b(1 + e^{cx})$, where a stands for the intercept, b for overall change in the parameter during the process, and c the rate of change), the only that partially fitted the pH changes. The overall comparison among treatments was also achieved by analyzing the areas below the curves of the several fermentation parameters vs. time (Bautista-Gallego et al., 2010). Their values were estimated by integration using the SigmaPlot v. 11 software package (Systat Software, Inc.).

The FDA approach used in this work consisted of transforming the data into functional objects and smooth curves by using a smoothing spline estimator (S_λ) (a cubic spline with continuous first two derivatives) based on the minimization of:

$$\sum_{i=1}^{n} (y_i - s(t_i))^2 - \lambda \int s''(t)^2 dt$$

where s stands for the second derivative of spline function and λ (the smoothing parameter) is intended for maintaining a balance between closeness of fit to the data (first term) and the roughness (sudden local variation) penalty (second term). The presentation of the full data set is usually visualized by plotting them vs. time. This functional object may be used for a series of estimations (pointwise evaluation, maximum, minimum, area below the curves, or an average of treatments and their standard deviation) as well as for the calculus of their first and second derivatives (the speed and acceleration, respectively, at which the changes occur).

Another approach for transforming data into smooth curves is by regression that is a type of functional linear model where the independent variables are indicator variables conveying membership in a combination of factor levels:

$$y_i(t) = \beta_0(t) + \sum_{i=1}^{q} x_{ij}\beta_j(t) + \varepsilon_i(t)$$

where $y_i(t)$ is a functional response, the values of x_{ij} are either 0 or 1, $i = 1, 2, \ldots$ N for curves, $j = 1, 2 \ldots .q$ for groups or products and $\sum_{i=1}^{q} \beta_j(t) = 0$ for all t. As result of its application, the corresponding fitted curves are obtained as well as maxima, minimum, and derivatives. These results may be also used for the comparison of treatments by fANOVA, using permutation testing, which is performed by permutation of observations across groups. The percentage of repetitions in which the calculated values of F exceeded the Fs obtained from the original data, is the p-value under de null hypothesis. The output of the test includes the fitted curves, their maximum (and its corresponding time), the overall p-value of the permutation F-test, a plot of the observed pointwise statistic values and the 0.05 critical values. A functional permutation t-test, performed similarly to the F-test permits the pointwise comparison between two treatments. The FDA was achieved using the R routines and "fda" functions for R software v. 3.2.1 (https://www.r-project.org/) developed by Bi and Kuesten (2013). Therefore, those interested in FDA application are kindly referred to their R routines and tutorial.

The relationships among the several outputs obtained from PDO *Aloreña de Málaga* fermentation, using both the conventional and the FDA approaches, were studied by multivariate techniques, particularly the exploratory biplot (Gabriel, 1971), which allows displaying