



タイトル Title	Monitoring of Water Spectral Pattern of lactobacilli development as a tool for rapid selection of probiotic candidates
著者 Author(s)	Slavchev, Aleksandar / Kovacs, Zoltan / Koshiba, Haruki / Bazar, Gyorgy / Pollner, Bernhard / Krastanov, Albert / Tsenkova, Roumiana
掲載誌・巻号・ページ Citation	Journal of Near Infrared Spectroscopy,25(6):423-431
刊行日 Issue date	2017-10-20
資源タイプ Resource Type	Journal Article / 学術雑誌論文
版区分 Resource Version	author
権利 Rights	Copyright © 2017 by SAGE Publications
DOI	10.1177/0967033517741133
JaLDOI	
URL	http://www.lib.kobe-u.ac.jp/handle_kernel/90006061



Monitoring of Water Spectral Patterns of Lactobacilli Development as a Tool for Rapid Selection of Probiotic Candidates

Journal:	<i>Journal of Near Infrared Spectroscopy</i>
Manuscript ID	JNS-17-0025.R2
Manuscript Type:	Original Research Article
Date Submitted by the Author:	14-Oct-2017
Complete List of Authors:	<p>Slavchev, Aleksandar; University of Food Technologies, Microbiology; Kobe University, Graduate School of Agricultural Science, Biomeasurement Technology Laboratory</p> <p>Kovács, Zoltán ; Szent Istvan Egyetem, Department of Physics and Control; Kobe University, Graduate School of Agricultural Science, Biomeasurement Technology Laboratory</p> <p>Koshiba, Haruki; Kobe University, Graduate School of Agricultural Science, Biomeasurement Technology Laboratory</p> <p>Bazar, Gyorgy; Kaposvár University, Faculty of Agricultural and Environmental Sciences, Department of Nutritional Science and Production Technology; Kobe University, Graduate School of Agricultural Science, Biomeasurement Technology Laboratory</p> <p>Pollner, Bernhard; Medical University of Innsbruck, Department for Hygiene and Medical Microbiology; Kobe University, Graduate School of Agricultural Science, Biomeasurement Technology Laboratory</p> <p>Krastanov, Albert; University of Food Technologies, Biotechnology</p> <p>tsenkova, roumiana; Kobe University, Graduate School of Agricultural Science, Biomeasurement Technology Laboratory</p>
Keywords:	probiotics, aquaphotomics, near infrared spectroscopy, NIR, chemometrics, lactic acid bacteria
Abstract:	<p>Nowadays a quick and inexpensive method, which allows rapid, in vivo comprehensive probiotic bacteria identification, is needed. To elucidate a new concept to evaluate probiotic bacteria, NIR spectroscopy with aquaphotomics were applied to monitor the growth of eight <i>Lactobacillus bulgaricus</i> and one <i>Lactobacillus gasseri</i> bacteria strains. Their resistance to low pH (1.8) in the presence of pepsin and bile were measured and further used as reference data for analysis of the simultaneously acquired spectral data. The acquired spectral data in the region of 1100-1300 nm were subjected to various methods for multivariate data analyses – PCA, LDA, SIMCA and PLSR. The results showed high accuracy of bacteria strains classification according to their resistance and the potential of the tested wavelength region for rapid selection and prediction of some basic phenotypic characteristics of probiotic candidates. Results of the current study also revealed different suitability of each growth stage when using NIR spectra for the classification of bacteria strains.</p>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

SCHOLARONE™
Manuscripts

For Peer Review

1
2
3
4 1 Monitoring of Water Spectral Patterns of Lactobacilli Development as a Tool for
5
6 2 Rapid Selection of Probiotic Candidates
7
8
9
10 3

11
12 4 Aleksandar Slavchev^{1,2}, Zoltan Kovacs^{1,3}, Haruki Koshiba¹, Gyorgy Bazar^{1,4}, Bernhard Pollner^{1,5},
13
14
15 5 Albert Krastanov^{6,7}, Roumiana Tsenkova^{1*}
16

17
18 6 ¹Kobe University, Graduate School of Agricultural Science, Biomeasurement Technology Laboratory,
19
20
21 7 Kobe, Japan
22

23
24 8 ²University of Food Technologies, Department of Microbiology, Plovdiv, Bulgaria
25

26
27 9 ³Szent István University, Faculty of Food Science, Department of Physics and Control, Budapest,
28
29
30 10 Hungary
31

32
33 11 ⁴Kaposvár University, Department of Nutritional Science and Production Technology, Faculty of
34
35
36 12 Agricultural and Environmental Sciences, Kaposvár, Hungary
37

38
39 13 ⁵Department for Hygiene and Medical Microbiology, Medical University of Innsbruck, Innsbruck,
40
41
42 14 Austria
43

44
45 15 ⁶University of Food Technologies, Department Biotechnology, Plovdiv, Bulgaria
46

47
48 16 ⁷Bulgarian Academy of Science, Specialized Academic Unit for Traditional Bulgarian Dairy
49
50
51 17 Products, Plovdiv, Bulgaria
52

53
54
55
56 18
57 19 Corresponding author: prof. Roumiana Tsenkova, Kobe University, Faculty of Agriculture,
58
59
60

1
2
3
4 1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Kobe, Japan, Tel & Fax: +81 78 803 5911,

5
6
7 2 e-mail: rtsen@kobe-u.ac.jp
8
9

10 3
11

12 4 **Abstract**

13
14
15 5 Nowadays a quick and inexpensive method, which allows rapid, *in vivo* comprehensive probiotic
16
17
18 6 bacteria identification, is needed. To elucidate a new concept to evaluate probiotic bacteria, NIR
19
20
21 7 spectroscopy with aquaphotomics were applied to monitor the growth of eight *Lactobacillus*
22
23
24 8 *bulgaricus* and one *Lactobacillus gasseri* bacteria strains. Their resistance to low pH (1.8) in the
25
26
27 9 presence of pepsin and bile were measured and further used as reference data for analysis of the
28
29
30 10 simultaneously acquired spectral data. The acquired spectral data in the region of 1100-1300 nm
31
32
33 11 were subjected to various methods for multivariate data analyses – PCA, LDA, SIMCA and PLSR.
34
35
36 12 The results showed high accuracy of bacteria strains classification according to their resistance and
37
38
39 13 the potential of the tested wavelength region for rapid selection and prediction of some basic
40
41
42 14 phenotypic characteristics of probiotic candidates. Results of the current study also revealed different
43
44
45 15 suitability of each growth stage when using NIR spectra for the classification of bacteria strains.
46
47
48 16

49
50 17 Keywords: probiotics, aquaphotomics, near infrared spectroscopy, NIR, chemometrics, lactic acid
51
52
53 18 bacteria.
54

55 56 19 **Introduction** 57 58 59 60

1
2
3
4 1 Lactic acid bacteria (LABs) are a diverse group of gram-positive, non-sporulent bacteria, anaerobes or
5
6
7 2 facultative anaerobes requiring specific nutrition environment, and whose ultimate metabolite product
8
9
10 3 is lactic acid. Currently, LABs became microorganisms of industrial importance for their fermentation
11
12
13 4 activity, as well as for their health and nutrition benefits. LABs, particularly lactic bacteria, which
14
15
16 5 occupy important niches in the gastro-intestinal tract (GIT) are considered as offering numerous
17
18
19 6 probiotic benefits to general health status and proper physical condition.

20
21 7 Probiotics are live microorganisms which when administered in adequate amounts confer a health
22
23
24 8 benefit on the host. They express their positive effect by balancing of the gastro-intestinal tract
25
26
27 9 microflora, improve the immune response ¹, decrease the symptoms of lactose intolerance and
28
29
30 10 allergies in susceptible individuals ², reduce the risk of cancer ³, and alleviate irritable bowel
31
32
33 11 syndrome and inflammatory bowel diseases ^{4,5}. The mechanism of probiotic activity has not been
34
35
36 12 established yet, but it probably includes modification of pH levels of the GIT ⁶, pathogens
37
38
39 13 antagonism through the production of antimicrobials ⁷, competition for receptor sites ⁸, nutrients and
40
41
42 14 growth factors, stimulation of immune cells ⁹ and lactase production ^{10,11}.

43
44
45 15 The most important probiotic characteristic is the ability of surviving the harsh environment in the
46
47
48 16 upper gastro-intestinal tract. This is because during their passage through the stomach and the small
49
50
51 17 intestine they face an environment with low pH and high concentrations of gall ¹¹. In the genus
52
53
54 18 *Lactobacillus*, there are strains which differ in resistance to the conditions in the upper
55
56
57 19 gastro-intestinal tract. Strains with high tolerance are of interest for the production of probiotics and
58
59
60

1
2
3
4 1 probiotic food. The main issue in this field is the selection of strains which exhibit strong probiotic
5
6
7 2 characteristics. Currently, two main strategies have been applied for the selection of probiotic strains:
8
9
10 3 selection of strains based on the presence of particular genes and *in vitro* examination of strain
11
12
13 4 resistance^{12,13}. These methods are time-consuming, require expensive equipment and consumables,
14
15
16 5 and they give uncertain results. Therefore, a quick and inexpensive method, which allows rapid, *in*
17
18
19 6 *vivo* comprehensive probiotic bacteria selection, is needed¹¹.

20
21 7 A new approach to this issue has been recently proposed. It is based on using near infrared (NIR)
22
23
24 8 spectroscopy and aquaphotomics for real-time monitoring of growing lactobacilli to classify and
25
26
27 9 discriminate strains possessing probiotic characteristics. Only by analysing their NIR spectra
28
29
30 10 obtained in the time period 11.4-12 h of the cultivation time¹¹ has it been possible to predict their
31
32
33 11 resistance to conditions similar to those in the upper gastro-intestinal tract.

34
35
36 12 The main goal of the study reported in this paper was to use data from the region between 1100 and
37
38
39 13 1300 nm to find the most suitable growth stage for *in vivo* spectral monitoring and discrimination of
40
41
42 14 resistant and non-resistant *Lactobacillus* strains. A further aim was to predict the bile minimal
43
44
45 15 inhibitory concentration and low pH resistance of *Lactobacillus* strains based on their NIR
46
47
48 16 spectra obtained during the cultivation process.

50 **Materials and Methods**

51 *Bacterial strains*

52
53
54
55
56 19 Five probiotic and four non-probiotic strains (genus *Lactobacillus*) possessing different bile salt
57
58
59
60

1 tolerance and ability to resist low pH (pH 1.80) in the presence of pepsin were used: probiotic strains
2
3
4 1 tolerance and ability to resist low pH (pH 1.80) in the presence of pepsin were used: probiotic strains
5
6
7 2 *L. bulgaricus* S06, *L. gasseri* S20 (provided by "Selur Pharma" Ltd. - Bulgaria), *L. bulgaricus* S22, *L.*
8
9
10 3 *bulgaricus* S11, *L. bulgaricus* S10 and non-probiotic strains - *L. bulgaricus* S09, *L. bulgaricus* Y12,
11
12
13 4 *L. bulgaricus* S03 and *L. bulgaricus* S02. The strains were provided by "Specialized Academic Unit
14
15
16 5 for Traditional Bulgarian Dairy Products within the Bulgarian Academy of Sciences" with exception
17
18
19 6 of *L. bulgaricus* Y12 which was isolated from yoghurt. All microorganisms were freeze-dried and
20
21
22 7 kept at -80 °C.

23 24 8 *Preparation of stock cultures and active bacterial culture*

25
26
27 9 The strains were cultivated in MRS broth (De Man, Rogosa and Sharpe agar; Merck, Tokyo, Japan)
28
29
30 10 at 37 °C for 24 h. The biomass obtained after centrifugation at 5000 min⁻¹ for 5 min was twice
31
32
33 11 washed with PBS (Phosphate-buffered saline) buffer (pH 7.00) and suspended in 15 %w/v glycerol
34
35
36 12 solution to the initial volume and stored at -80 °C for further use.

37
38
39 13 Tubes containing 1 ml MRS broth were inoculated with 50 µl glycerol suspension of stock culture
40
41
42 14 and cultivated for 18 – 20 h at 37 °C.

43 44 15 *Determination of the optical density of the bacterial cultures*

45
46
47 16 The optical density was determined by using a micro-plate reader iMark (BioRad, Hercules,
48
49
50 17 California, USA) against MRS broth as blank at $\lambda = 665$ nm. The sample volume was 150 µl with
51
52
53 18 correction to 1 cm path length. Every optical density is presented as an average value of nine optical
54
55
56 19 densities obtained from three independent samples (tubes or deep well plate wells) measured three
57
58
59

1
2
3
4 1 consecutive times.
5
6

7 2
8
9

10 3 *Determination of the resistance to low pH value in presence of pepsin*
11

12 4 A modified method of Pitino ¹³ was used. 750 µl MRS-broth were inoculated with 50 µl of active
13
14
15 5 bacterial culture and cultivated for 18 – 20 h at 37 °C. The culture medium was then centrifuged at
16
17
18 6 10000 min⁻¹ for 5 min, the biomass was washed twice with PBS buffer (pH 7.00). The cells were
19
20
21 7 suspended to the original volume with the low pH buffer (pH 1.8), containing HCl (0.2 M), NaCl
22
23
24 8 (0.08 M), CaCl₂ (0.03 mM), and pepsin from porcine gastric mucosa (9000 U ml⁻¹; Wako, Japan).
25
26

27 9 After 3 h cultivation at 37 °C the biomass was centrifuged at 10000 min⁻¹ for 5 min and washed with
28
29
30 10 PBS buffer and re-suspended to its original volume with PBS buffer. Tubes containing 750 µl MRS
31
32
33 11 broth were inoculated with 50 µl of low-pH treated cells suspensions and cultivated at 37 °C for 24 h.
34
35
36 12 The optical density of the cultures was measured at λ=665 nm at 0h and 24 h. The resistance each
37
38
39 13 strain to low pH in the presence of pepsin was evaluated by cell growth and is presented by the
40
41
42 14 increase in the optical density of the culture medium after 24h cultivation at 37 °C (Yield of biomass
43
44
45 15 after 3 h stay at pH 1.8 and 9000 U ml⁻¹ pepsin).
46
47

48 16 *Determination of bile minimal inhibitory concentration*
49

50 17 MRS broths (750 µl) with double-fold decreasing concentrations of dry bile (Wako, Japan)
51
52
53 18 0.156-5.000 mg/ml were inoculated with 50 µl active bacterial culture and cultivated at 37 °C for 24
54
55
56 19 h and the optical density at λ=665 nm was determined.
57
58
59
60

1
2
3
4 1 *Monitoring of the cultures by NIR Spectroscopy*

5
6
7 2 MRS broth (15 ml) was inoculated with about 0.5 ml active bacterial culture to OD=0.1 ($\lambda=665$) and
8
9
10 3 cultivated at 37 °C for 24 h with shaking on a vibratory shaker in a 50 ml centrifuge tube. The NIR
11
12
13 4 transfectance spectra of the culture were acquired in the entire spectral region (400-2500 nm) with
14
15
16 5 0.5 nm step (4200 data points) at every 4 min by using a FOSS XDS OptiProbe Analyzer attached
17
18
19 6 with immersion type probe (FOSS NIRSystems, Inc., Hoganas, Sweden or Hilleroed, Denmark,
20
21
22 7 recently distributed by Metrohm NIRSystems AG, Herisau, Switzerland). A reference spectrum was
23
24
25 8 taken at the beginning of every measurement series by placing the immersion probe in the dark
26
27
28 9 aperture position of the instrument. The spectra taken in the first 40 min of the cultivation time were
29
30
31 10 discarded and those after 40 min until the scan of 20 h of the monitoring were used for data
32
33
34 11 evaluation. Spectra acquisition was performed with the VISON 3.50 (FOSS NIRSystems, Inc.,
35
36
37 12 Hoganas, Sweden) software. After pre-experiments, 0.5 mm layer thickness (set by spacer) was
38
39
40 13 found to be the most appropriate to achieve an applicable signal in the first overtone region of water.

41 14 *Spectral Data Analysis*

42
43
44 15 The water first overtone of combination vibrations at the wavelength range 1100-1300 nm was used
45
46
47 16 for data evaluations. As a first step of spectral pretreatment, smoothing by using a Savitzky-Golay¹⁴
48
49
50 17 filter with 21 data points and second polynomial order was applied. For eliminating the scattering
51
52
53 18 effect MSC (multiplicative scatter correction) transformation¹⁵ was performed.

54
55
56 19 Principal Component Analysis¹⁶ (PCA) was used to discover the multidimensional pattern of
57
58
59
60

1
2
3
4 1 variations in the strains low pH and bile resistance scores and the NIR spectral dataset obtained at
5
6
7 2 whole cultivation time (40-1200 min) and to provide orthogonal variables for supervised
8
9
10 3 classifications of linear discriminant analysis using the spectral dataset of different growth stages.

11
12 4 Linear Discriminant Analysis (LDA)¹⁷ was performed to generate supervised qualitative models to
13
14
15 5 identify the groups of bacterial strains. The scores of the first ten principal components (PC) of the
16
17
18 6 PCA models were used in LDA as input variables to build models for the classification of the three
19
20
21 7 main groups of the bacteria strains. LDA models were tested with “one-strain-out” cross-validation
22
23
24 8 (described in more details below) and with independent prediction using the spectra of *L. bulgaricus*
25
26
27 9 S22, *L. bulgaricus* S10 and *L. bulgaricus* Y12.

28
29
30 10 To find relationship between spectral data and phenotype parameters of bacterial strains (the
31
32
33 11 minimum inhibitory concentration of bile(MIC) and ability to recover after 3 h stay at low pH) we
34
35
36 12 applied Partial Least Squares Regression¹⁵ (PLSR). The PLSR models were evaluated by the
37
38
39 13 coefficient of determination in cross-validation (R^2_{cv}) and standard error of cross-validation (SECV).
40
41
42 14 The maximum number of latent variables was determined as $1/10^{\text{th}}$ of the number of observation (n)
43
44
45 15 to avoid overfitting. The PLSR models were validated using the “one-strain-out” validation method.
46
47
48 16 The data set was split into training and test sets. The spectral data of 8 strains were used as the
49
50
51 17 training set; and those of one strain left, as the test set. This process of data splitting was repeated 9
52
53
54 18 times to ensure that the data of all the strains was included in the evaluation set once¹⁷.

55
56 19 For qualitative assessment of strains for low-pH and bile tolerance Soft Independent Modeling of
57
58
59
60

1
2
3
4 1 Class Analogy (SIMCA) analysis was used. Raw spectra were mean centered. The cross-validation
5
6
7 2 was performed using “one-strain-out” in the same way as is described above.
8

9
10 3 Aquagrams¹⁸ were calculated in order to show the differences of the absorbance values at the water
11
12 4 matrix coordinates (WAMACs) for the group of resistant and non-resistant bacteria. The star-chart
13
14
15 5 displays averaged normalized spectral absorbance values of the groups of resistant and non-resistant
16
17
18 6 bacteria (acquired at the same conditions) cultivated in MRS broth. The spectra of each strain were
19
20
21 7 acquired at 37 °C in the time intervals of 0.3-6.3 h and 6.3-12.3 h of the cultivation time (phases III
22
23
24 8 and IV).
25

26
27 9 The scripts for PLSR, LDA and aquagram calculation and visualization were written and executed in
28
29
30 10 R-project environment (RStudio Ver. 0.98 and R Ver. 3.0.1, R Foundation for Statistical Computing,
31
32
33 11 Vienna, Austria). The calculation and visualization of PCA and SIMCA were performed with
34
35
36 12 SimcaVer. 13.5 (Umetrics AB, USA; <http://umetrics.com/>) and Piroette 4 (Infometrix, USA;
37
38
39 13 <https://infometrix.com/software/>).
40

41
42 14

43 44 15 **Results and Discussions**

45
46
47 16 *Classification of the strains based on their bile minimal inhibitory concentration (MIC) and low pH*
48
49
50 17 *resistance*
51

52
53 18 Based on the low pH tolerance (Yield of biomass after 3 h stay at pH 1.8 and 9000 U ml⁻¹pepsin) and
54
55
56 19 the MIC for bile, all strains were divided in two groups – strains which were capable of growth after
57
58

1
2
3
4 1 3h stay at pH 1.8 in the presence of 9000 U ml⁻¹ pepsin and could survive and grow at high bile
5
6
7 2 concentrations and strains which were not resistant to that environment. *L. bulgaricus* S06 showed
8
9
10 3 medium results and was not included in either of the two groups (Table 1).
11

12
13 4 Table 1.

14
15 5 Strains varied in their resistance to low pH in the presence of pepsin and bile. These are essential
16
17
18 6 characteristics, which allow strains to survive in the human gastro-intestinal tract. This is due to their
19
20
21 7 adaptation ability and depends on the presence and the expression of some genes, which leads to
22
23
24 8 differences in the levels of some proteins and enzyme activities ¹⁹.

25
26
27 9 PCA results based on the reference data i.e. bile tolerance and pH resistance for each strain are
28
29
30 10 shown in Fig. 1.
31

32
33 11 Figure 1.

34
35
36 12 The first principal component (PC1) of this matrix explained 98.3 % of the total variance. Its scores
37
38
39 13 were highly correlated with the probiotic properties of the strain and represented very well the ability
40
41
42 14 of the strains to grow in the presence of bile and to survive at very low pH. On the score plot (Fig. 1)
43
44
45 15 the group of resistant strains is very well separated from the group of non-resistant ones. Loadings of
46
47
48 16 the model (indicated by the arrows) show that both parameters were correlated.

49
50 17 *Determination of Bacterial Growth Phases*
51

52
53 18 The typical microbial batch culture growth consists of four growth phases: Lag phase, Exponential
54
55
56 19 phase, Stationary phase and Death phase ²⁰. They are not always clearly defined because of longer
57
58
59

1 transition periods which may occur due to different factors. In order to separate the growing curves
2 into growth phases a modified model of Gompertz and its fourth derivative were applied^{21,22} (Fig.
3 2).

$$4 \quad y = A \cdot \exp \left\{ -\exp \left[\frac{\mu_m e}{A} (\lambda - t) + 1 \right] \right\},$$

5 where A is the maximal optical density, μ_m is the maximal specific growth rate, λ is the length of the
6 lag-phase and t is the time.

7 Figure 2.

8 The fourth derivative of the absorbance at 600 nm changes in the model was used to divide all
9 growth curves into five phases (Fig. 2), which is in accordance with Widdel's model of typical
10 growing bacterial culture²³. The accurate separation of bacterial growth cycle is essential for further
11 data analyses and allows us to compare strains at their respective growth stage and at equal
12 physiological condition.

13 *Spectral data analyses*

14 Near infrared spectra of the growing bacterial cultures changed continuously during the whole
15 cultivation period. In our previous work¹¹ we showed that the region of 1100 – 1850 nm can be used
16 successfully to classify probiotic lactobacilli and predict some of their essential properties at certain
17 periods during the cultivation period (11.4 – 12h) where the cultures are present in large numbers of
18 similar cells, which are still viable. We have found that water absorbance bands in the first overtone
19 region were the most informative for classification. Here, we are exploring the first overtone of the

1 combination vibrations of water, in the range 1100 – 1300 nm Figure 3), which has less absorbance
2 and, therefore, it is most suitable when using longer path lengths.

3 Figure 3.

4 Raw spectra were MSC transformed to eliminate the scattering effect and baseline differences (Fig.
5 3), which makes them more suitable for further data analyses. The cultivation time was separated
6 into growing phases by using 4th derivative of the modified Gompertz model and every time period
7 was used to build LDA, SIMCA and PLS regression models.

8 Principal component analysis of spectral data acquired in the time period of 40-1200 min of
9 cultivation showed that 99.95% of the whole variance was explained by PC1 (Fig. 4a). This variation
10 was observed due to the spectral changes that occurred during the growing time, but it also showed
11 the differences between the groups of resistant and non-resistant strains. All resistant strains were
12 grouped in the right part of the score plot, while the non-resistant ones in the right side this
13 separation were more prominent after Phase II.

14 Figure 4.

15 The loading vector (Fig. 4b) of PC1 shows peaks at **1120**, **1158** and **1283**nm. These wavelengths are
16 the most important for the separation of the strains and changed the most during the cultivation and
17 show differences in free and protonated water molecules²⁴.

18 The spectral data for strains were analyzed by Linear discriminant analysis and Soft Independent
19 Modeling of Class Analogy to test its potential for qualitative assessment of low-pH and bile

1 tolerance of strains.

2 The results of the classification matrix showed highest accuracy of prediction at earlier exponential
3 phase (phase III and IV) for both LDA and SIMCA methods (Table 2.).

4 Best classification of the resistant and non-resistant groups was found in phase IV with LDA where
5 the recognition ability was 98.6% and prediction ability was 100% for both cross-validation and
6 independent prediction (Fig. 4c). However, LDA models built on data of phase III also showed 100%
7 correct prediction ability (Table 2.).

8 Table 2.

9 Almost all of the spectra were classified correctly at the second and the third growth phases – 97.7%
10 and 96.5% respectively with SIMCA.

11 The discriminating power of the SIMCA model (Fig. 4d) built on the spectral data of phase III
12 showed peaks at 1116 nm, **1123** nm, 1147.5 nm, **1158.5** nm, 1184 nm, **1197** nm, 1225 nm, 1264 nm
13 and **1285** nm. Both PCA and SIMCA models showed consistency in the wavelengths with highest
14 importance for the discrimination of resistant and non-resistant *Lactobacillus* strains.

15 Partial least square regression models (PLSR) were built to determine relationship between spectral
16 data and the tolerance of strains to low pH and bile. Two models were built – one for prediction of
17 the ability of a strain to survive during 3h stay at pH 1.8 and 9000 U ml⁻¹ pepsin (Fig. 5a and Fig. 5c)
18 and another one for the bile minimal inhibitory concentration for a strain (Fig. 5b). The accuracy of
19 the models was evaluated by their R² of cross validation and SECV (standard error of

1
2
3
4 1 cross-validation).

5
6
7 2 Figure 5.

8
9
10 3 For the prediction of low pH resistance, the best results were observed at phases IV and V, while for
11
12 4 the prediction of bile MIC the most suitable phases were III and IV. The results of PLSR regression
13
14
15 5 models showed close correlation and relatively low error of cross-validation. In both cases models
16
17
18 6 using spectra from phase IV showed the highest accuracy, which confirms our previously obtained
19
20
21 7 results ¹¹.

22
23
24 8 The wavelengths which were important for the prediction of the low pH resistance of a strain are
25
26
27 9 shown on Fig. 5c. The regression vector showed two groups of wavelengths in the whole region of
28
29
30 10 1100-1300 nm. The first group repeats wavelengths already discovered by SIMCA modelling (Fig.
31
32
33 11 4d) wavelengths – 1116, 1148.5, 1183, 1199 and 1225 nm, while the other wavelengths were unique
34
35
36 12 for the regression model – 1134, 1154, 1176, 1190, 1234, 1241, 1255 and 1295 nm.

37
38
39 13 Another successfully applied aquaphotomics tool, aquagram, was used to examine the water
40
41
42 14 combination band region. It proposes seven specific spectral ranges which are of importance for
43
44
45 15 bacteria characterization. It is a star-chart which contains averaged normalized absorbance values at
46
47
48 16 wavelengths in those regions of interest. These values contain information about water molecular
49
50
51 17 conformations and their respective hydrogen and covalent bonds ¹¹. In this study we applied
52
53
54 18 aquagrams to the shorter wavelength region of 1100-1300 nm in order to show the differences
55
56
57 19 between non-resistant and resistant strains.

Figure 6.

The aquagrams of the monitored nine *Lactobacillus* strains calculated using the spectral data acquired in Phase III and Phase IV (Fig. 2) are presented on Figure 6. In both phases the group of resistant strains showed higher values at shorter wavelengths, i.e. in the range between 1118.5 and 1198.5 nm and, in contrast, the absorbance values at 1217.5 and 1224.5 nm were higher in the case of *Lactobacillus* S06 and the non-resistant strains.

It is interesting to note that although in both resistant and non-resistant groups we had four different strains the spectral patterns of the strains of the same type were similar while those of the different types are more dissimilar.

The pattern of the moderate strain showed more similarity to the pattern of the non-resistant group, but some differences were observed. In phase III the value of *L. bulgaricus* S06 was higher than the values of the non-resistant group at 1118.5 nm thus it had more similarities with the resistant one. However, in phase IV the overall pattern of *L. bulgaricus* S06 was similar to the one of the non-resistant group but showed higher absorption values between 1113.5 and 1198.5 nm.

The wavelengths presented in the aquagram are characteristic of protonated water molecules presented by the asymmetric OH-stretch vibrational frequencies of $[H^+ \cdot (H_2O)_{4,5}] - H_3O^+$ symmetric stretch and asymmetric stretch of 2nd overtone (at 1160 and 1168 nm respectively)²⁴. The band around 1217 nm could refer to 2nd overtone of water²⁵.

The selection of strains possessing high resistance to conditions similar to those in the human

1
2
3
4 1 digestive tract has been studied by many authors using different approaches. The main shortcomings
5
6
7 2 of the classical selection methodology were overcome by using different *omics* approaches, such as
8
9
10 3 proteomics, genomics etc. Recently we showed that the aquaphotomics, which is directly related to
11
12
13 4 system functionality, is a powerful tool and can be used successfully for rapid selection of probiotic
14
15
16 5 lactobacilli ¹¹. Its non-invasive methodology allows rapid and comprehensive evaluation of potential
17
18
19 6 probiotic candidates and might lead to understanding of new phenomena in microbiology.
20
21
22
23
24

25 8 **Conclusions**

26
27 9 With this work we build on the previous studies by analyzing the whole *Lactobacillus* cultivation
28
29
30 10 time period (divided into five stages) while having the same goal – to obtain high accuracy of
31
32
33 11 bacteria characterization and to find those characteristic wavelengths which can be used for rapid
34
35
36 12 classification of resistant strains and to assess their resistance to low pH and bile.

37
38
39 13 We found out that the first overtone of the water combination band in the shorter wavelength region
40
41
42 14 (1100-1300 nm) contains useful information about the resistance of strains, presented by differences
43
44
45 15 in the free water molecules, asymmetric O-H stretches and protonated water bands. These differences
46
47
48 16 are due to the presence of different cell metabolites and other chemical substances that change the
49
50
51 17 water molecular structure analyzed in this study and refer to the phenotype of every group of strains.

52
53 18 Highest accuracy of classification and calibration performance of the PLSR models for prediction of
54
55
56 19 strains resistant to low pH in presence of pepsin and to a high concentration of bile were observed in
57
58
59
60

1
2
3
4 1 the IV growth stage (6.3-12.3h of the cultivation time). This is in agreement with our previous
5
6
7 2 publication ¹¹, but also shows that it is possible to build PLSR models with high R² and low SECV
8
9
10 3 using spectra obtained at an even earlier time during the cultivation of *Lactobacillus*.

11 **Acknowledgments**

12
13
14
15
16 5 This work was supported by Japan Society for the Promotion of Science (JSPS), Grant Number
17
18
19 6 L14562 (Aleksandar Slavchev) and by the Janos Bolyai Research Scholarship of the Hungarian
20
21
22 7 Academy of Sciences and by the ÚNKP-17-4 New National Excellence Program of The Ministry of
23
24
25 8 Human Capacities (Zoltan Kovacs). The authors would like to thank Selur Pharma Ltd, Bulgaria for
26
27
28 9 providing two of the bacterial strains.
29
30
31
32

33 **References**

- 34
35
36 12 1. Parvez S, Malik KA, Ah Kang S, et al. Probiotics and their fermented food products are
37
38
39 13 beneficial for health. *J Appl Microbiol* 2006; 100: 1171–85.
40
41
42 14 2. Kalliomäki MA, Isolauri E. Probiotics and down-regulation of the allergic response. *Immunol*
43
44
45 15 *Allergy Clin North Am* 2004; 24: 739–52, viii.
46
47
48 16 3. Isolauri E. Dietary modification of atopic disease: Use of probiotics in the prevention of atopic
49
50
51 17 dermatitis. *Curr Allergy Asthma Rep* 2004; 4: 270–5.
52
53
54 18 4. McFarland L V. Meta-analysis of probiotics for the prevention of traveler's diarrhea. *Travel*
55
56 19 *Med Infect Dis* 2007; 5: 97–105.
57
58
59
60

- 1
2
3
4 1 5. Rastall RA, Gibson GR, Gill HS, et al. Modulation of the microbial ecology of the human
5
6
7 2 colon by probiotics, prebiotics and synbiotics to enhance human health: an overview of
8
9
10 3 enabling science and potential applications. *FEMS Microbiol Ecol* 2005; 52: 145–52.
- 11
12
13 4 6. Mack DR, Michail S, Wei S, et al. Probiotics inhibit enteropathogenic E. coli adherence in
14
15
16 5 vitro by inducing intestinal mucin gene expression. *Am J Physiol Gastrointest Liver Physiol*
17
18
19 6 1999; 276: G941-950.
- 20
21
22 7 7. Miraglia del Giudice M, De Luca MG. The role of probiotics in the clinical management of
23
24
25 8 food allergy and atopic dermatitis. *J Clin Gastroenterol* 2004; 38: S84-5.
- 26
27
28 9 8. Walker WA. Role of nutrients and bacterial colonization in the development of intestinal host
29
30
31 10 defense. *J Pediatr Gastroenterol Nutr* 2000; 30 Suppl 2: S2-7.
- 32
33
34 11 9. Pessi T, Sutas Y, Hurme M, et al. Interleukin-10 generation in atopic children following oral
35
36
37 12 *Lactobacillus rhamnosus* GG. *Clin Exp Allergy* 2000; 30: 1804–1808.
- 38
39
40 13 10. Marteau P, Flourie B, Pochart P, et al. Effect of the microbial lactase (EC 3.2.1.23) activity in
41
42
43 14 yoghurt on the intestinal absorption of lactose: An in vivo study in lactase-deficient humans.
44
45
46 15 *Br J Nutr* 2007; 64: 71.
- 47
48
49 16 11. Slavchev A, Kovacs Z, Koshiha H, et al. Monitoring of Water Spectral Pattern Reveals
50
51
52 17 Differences in Probiotics Growth When Used for Rapid Bacteria Selection. *PLoS One* 2015;
53
54
55 18 10: e0130698.
- 56
57
58 19 12. Zotta T, Asterinou K, Rossano R, et al. Effect of inactivation of stress response regulators on
59
60

- 1
2
3
4 1 the growth and survival of *Streptococcus thermophilus* Sfi39. *Int J Food Microbiol* 2009; 129:
5
6
7 2 211–20.
8
9
10 3 13. Pitino I, Randazzo CL, Mandalari G, et al. Survival of *Lactobacillus rhamnosus* strains in the
11
12 upper gastrointestinal tract. *Food Microbiol* 2010; 27: 1121–7.
13
14
15 5 14. Savitzky A, Golay MJE. Smoothing and Differentiation of Data by Simplified Least Squares
16
17 Procedures. *Anal Chem* 1964; 36: 1627–1639.
18
19
20
21 7 15. Naes T, Isaksson T, Fearn T, et al. *A user friendly guide to multivariate calibration and*
22
23 *classification*. Chichester, UK: NIR Publications, 2002.
24
25
26
27 9 16. Cowe IA, McNicol JW. The use of principal components in the analysis of near-infrared
28
29 spectra. *Appl Spectrosc* 1985; 39: 257–266.
30
31
32
33 11 17. Berrueta LA, Alonso-Salces RM, Héberger K. Supervised pattern recognition in food analysis.
34
35 *J Chromatogr A* 2007; 1158: 196–214.
36
37
38
39 13 18. Kinoshita K, Miyazaki M, Morita H, et al. Spectral pattern of urinary water as a biomarker of
40
41 estrus in the giant panda. *Sci Rep* 2012; 2: 856.
42
43
44 15 19. De Angelis M, Gobbetti M. Environmental stress responses in *Lactobacillus*: A review.
45
46 *Proteomics* 2004; 4: 106–122.
47
48
49
50 17 20. Madigan M, Clark D, Stahl D, et al. *Brock Biology of Microorganisms 13th edition*. 2010.
51
52
53
54 18 21. Perni S, Andrew PW, Shama G. Estimating the maximum growth rate from microbial growth
55
56 curves: definition is everything. *Food Microbiol* 2005; 22: 491–495.
57
58
59
60

- 1
2
3
4 1 22. Mytilinaios I, Salih M, Schofield HK, et al. Growth curve prediction from optical density data.
5
6
7 2 *Int J Food Microbiol* 2012; 154: 169–76.
8
9
10 3 23. Widdel F. Theory and measurement of bacterial growth. *Di dalam Grundpraktikum Mikrobiol*
11
12 4 2007; 1–11.
13
14
15 5 24. Headrick JM, Diken EG, Walters RS, et al. Spectral signatures of hydrated proton vibrations
16
17 6 in water clusters. *Science (80-)* 2005; 308: 1765–1769.
18
19
20
21 7 25. Siesler HW, Ozaki Y, Kawata S, et al. (eds). *Wiley: Near-Infrared Spectroscopy: Principles,*
22
23 8 *Instruments, Applications.* Weinheim, Germany: Wiley-VCH Verlag GmbH, 2001. Epub ahead
24
25 9 of print November 2001. DOI: 10.1002/9783527612666.
26
27
28
29
30
31

32 Tables

33
34
35 **Table 1.** Bile minimal inhibitory concentration (MIC) of bile and yields of biomass after low pH
36 stress in presence of pepsin of the strains
37

38 Strains	39 Group	40 MIC of Bile, 41 mg/ml	42 Yield of biomass after 3 h 43 stay at pH 1.80 and 9000 44 U ml ⁻¹ pepsin (±SD*)
45 <i>L. bulgaricus</i> S10	46 Resistant	47 2.500	48 0.126±0.014
49 <i>L. bulgaricus</i> S11	50 Resistant	51 2.500	52 0.100±0.014
53 <i>L. gasseri</i> S20	54 Resistant	55 2.500	56 0.114±0.011
57 <i>L. bulgaricus</i> S22	58 Resistant	59 2.500	60 0.117±0.016
<i>L. bulgaricus</i> S06	Moderate	1.250	0.080±0.010
<i>L. bulgaricus</i> Y12	Non-resistant	0.625	0.041±0.006
<i>L. bulgaricus</i> S09	Non-resistant	0.156	0.029±0.005
<i>L. bulgaricus</i> S02	Non-resistant	0.156	0.006±0.001
<i>L. bulgaricus</i> S03	Non-resistant	0.313	0.005±0.002

* - Standard deviations calculated based on three parallel samples scanned three consecutive times

Table 2. Average recognition and prediction ability of SIMCA and LDA models built for the classification of resistant and non-resistant *Lactobacillus* strains derived from the spectroscopic data in the range between 1100 and 1300 nm

	SIMCA		LDA		
	model (%)	CV (%)	model (%)	CV (%)	prediction (%)
Phase II	79.77	77.06	96.79	94.87	93.16
Phase III	98.77	97.73	97.78	100	100
Phase IV	97.53	96.46	98.61	100	100
Phase V	81.13	78.85	98.92	98.29	99.71

CV – “one-strain-out” cross-validation

prediction – independent prediction using data of *L. bulgaricus* S22, *L. bulgaricus* S10 and *L. bulgaricus* Y12.

1 Captions to figures

2
3
4
5
6
7 **Figure 1.** PCA Bi-plot calculated when using only reference data (bile MIC and the yield of biomass
8
9
10 after 3h stay at pH 1.8 in presence of pepsin (9000 U ml⁻¹), reference data in Table 1) of the
11
12 *Lactobacillus* strains

13
14
15 **Figure 2.** Bacterial growth curve presented by OD at 600 nm and its 4th derivation of modified
16
17
18 Gompertz fitting (*L. bulgaricus* S06) showing the existence of five different growth stages

19
20
21 **Figure 3.** Truncated (1100 – 1300 nm) raw spectra (n=2610) and b) MSC transformed spectra of the
22
23
24 analyzed *Lactobacillus* strains, obtained in the time interval of 40 – 1200 min

25
26
27 **Figure 4.** PCA, LDA and SIMCA analyses on the spectral data of the *Lactobacillus* strains, using
28
29
30 1100-1300 nm wavelength interval. a) and b) PCA score plot (n=2610), and loading vectors of PCA
31
32
33 model calculated on data obtained in the time period of 40-1200 min, c) Scores of the LDA model
34
35
36 built for the classification of resistant and non-resistant *Lactobacillus* strains in growth Phase IV
37
38
39 (n_{model}=540, n_{prediction}=270) d) Discrimination power of SIMCA model built for the classification of
40
41
42 resistant and non-resistant *Lactobacillus* strains in growth Phase III (n=405).

43
44
45 **Figure 5.** PLSR models on spectral data obtained in different growth phases of the cultivation
46
47
48 process, wavelength interval 1100-1300 nm (n_{PhaseII}=351, n_{PhaseIII}=405, n_{PhaseIV}=810, n_{PhaseV}=1044).

49
50
51 Accuracy of prediction of a) strains resistant to low pH in presence of pepsin and b) strains MIC of
52
53
54 bile. c) Regression vector of the PLSR model built to predict strains resistant to low pH in presence
55
56
57 of pepsin in growth Phase IV.

1
2
3
4 **Figure 6.** Aquagrams on the spectra between 1100 and 1300 nm of *Lactobacillus* strains. Averaged
5
6
7 values of normalized absorbance values of the water matrix coordinates for every strain are plotted
8
9
10 on each axis. Results were calculated on spectral data obtained a) in growth Phase III (n=405) and b)
11
12
13 in growth Phase IV (n=810).
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

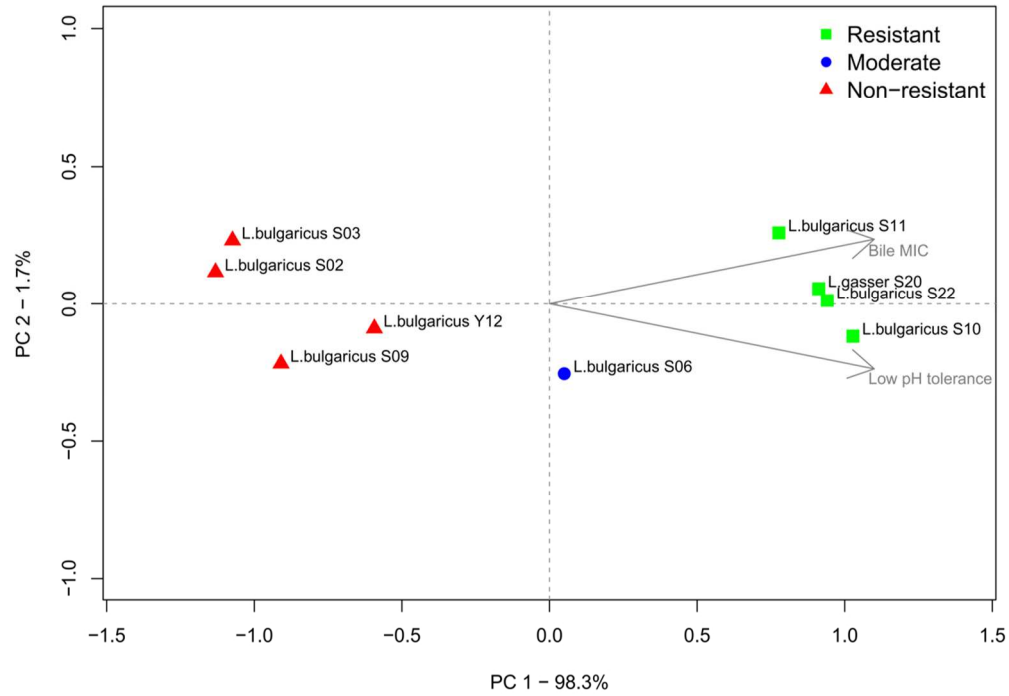


Figure 1. PCA Bi-plot calculated when using only reference data (bile MIC and the yield of biomass after 3h stay at pH 1.8 in presence of pepsin (9000 U ml⁻¹), reference data in Table 1) of the Lactobacillus strains

369x256mm (96 x 96 DPI)

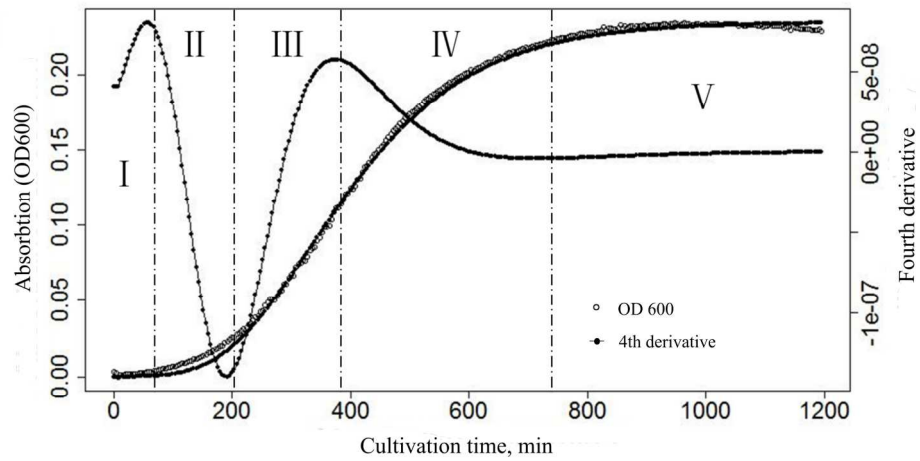


Figure 2. Bacterial growth curve presented by OD at 600 nm and its 4th derivation of modified Gompertz fitting (*L. bulgaricus* S06) showing the existence of five different growth stages

201x117mm (299 x 299 DPI)

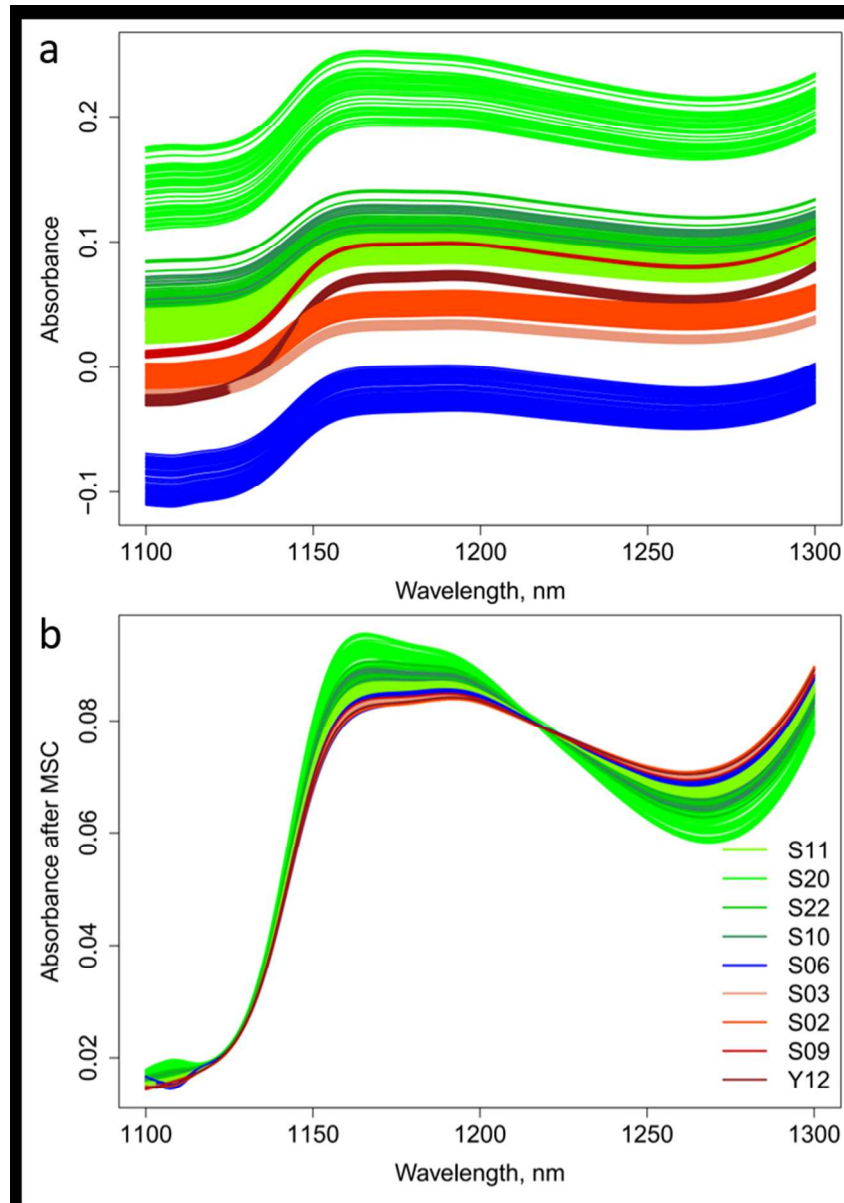


Figure 3. Truncated (1100 – 1300 nm) raw spectra (n=2610) and b) MSC transformed spectra of the analyzed *Lactobacillus* strains, obtained in the time interval of 40 – 1200 min

199x286mm (96 x 96 DPI)

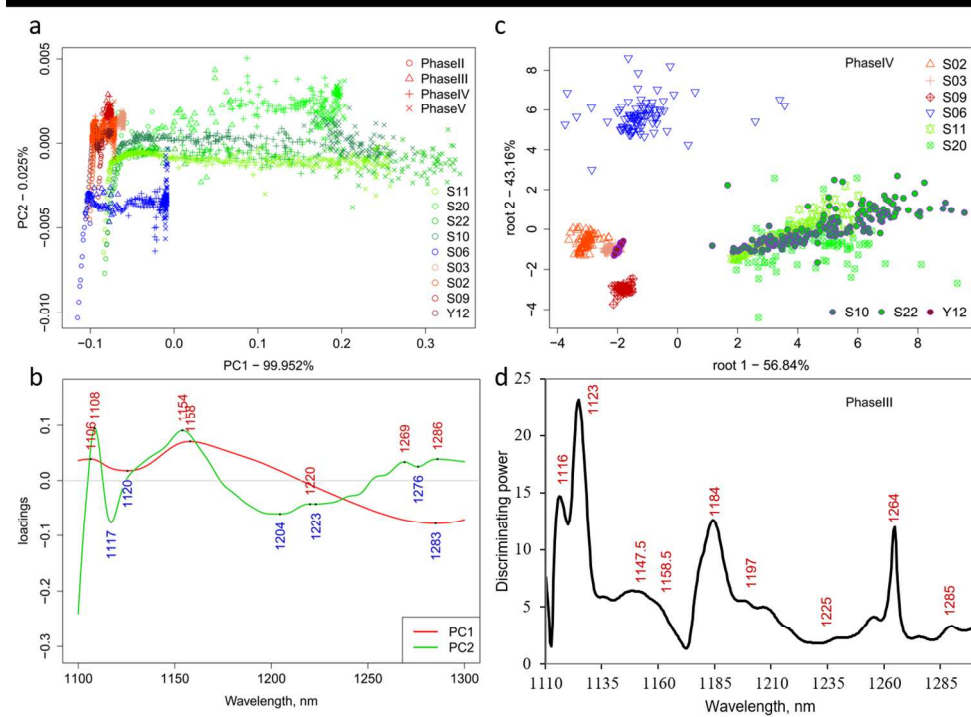


Figure 4. PCA, LDA and SIMCA analyses on the spectral data of the *Lactobacillus* strains, using 1100-1300 nm wavelength interval. a) and b) PCA score plot ($n=2610$), and loading vectors of PCA model calculated on data obtained in the time period of 40-1200 min, c) Scores of the LDA model built for the classification of resistant and non-resistant *Lactobacillus* strains in growth Phase IV ($n_{\text{model}}=540$, $n_{\text{prediction}}=270$) d) Discrimination power of SIMCA model built for the classification of resistant and non-resistant *Lactobacillus* strains in growth Phase III ($n=405$)

397x291mm (96 x 96 DPI)

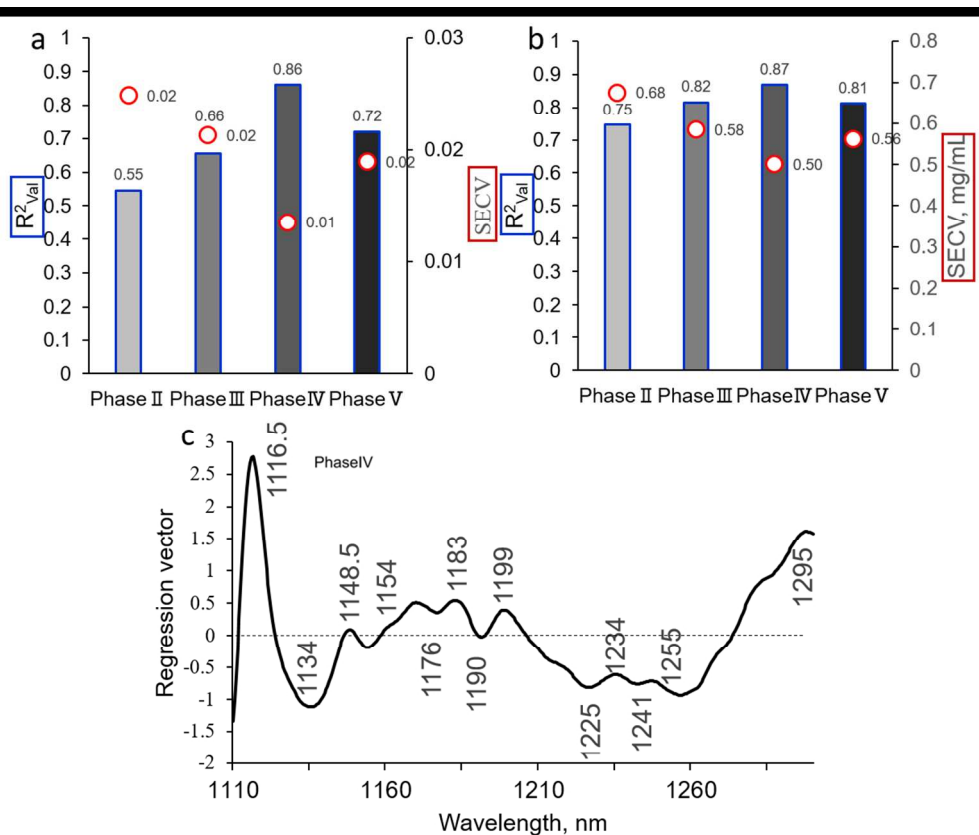


Figure 5. PLSR models on spectral data obtained in different growth phases of the cultivation process, wavelength interval 1100-1300 nm (nPhaseII=351, nPhaseIII=405, nPhaseIV=810, nPhaseV=1044). Accuracy of prediction of a) strains resistant to low pH in presence of pepsin and b) strains MIC of bile. c) Regression vector of the PLSR model built to predict strains resistant to low pH in presence of pepsin in growth Phase IV.

328x279mm (96 x 96 DPI)



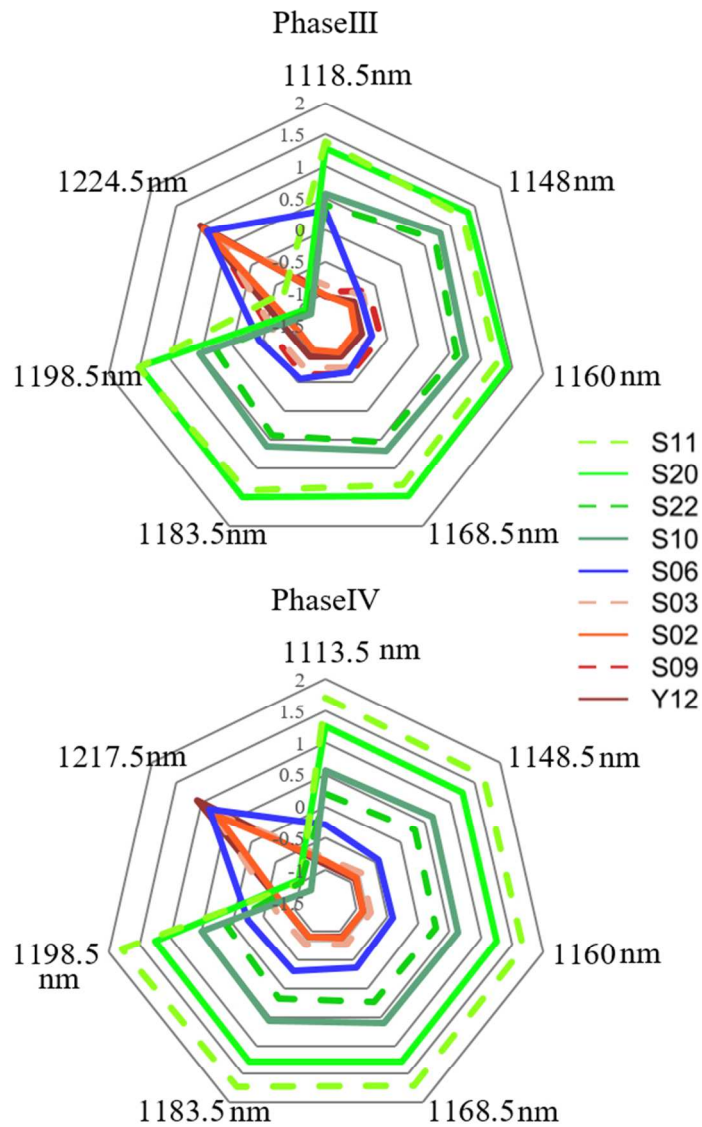


Figure 6. Aquagrams on the spectra between 1100 and 1300 nm of *Lactobacillus* strains. Averaged values of normalized absorbance values of the water matrix coordinates for every strain are plotted on each axis. Results were calculated on spectral data obtained a) in growth Phase III (n=405) and b) in growth Phase IV (n=810)

198x287mm (96 x 96 DPI)