ISSN 1330-9862
(FTB-2848)

*original scientific paper*

# Identification of Lactic Acid Bacteria and Propionic Acid Bacteria using FTIR Spectroscopy and Artificial Neural Networks

*Bartłomiej Dziuba* and Beata Nalepa*

University of Warmia and Mazury in Olsztyn, Faculty of Food Science, Department of Industrial and Food Microbiology, Plac Cieszyński 1, PL-10-957 Olsztyn, Poland

Received: June 27, 2011
Accepted: February 3, 2012

**Summary**

In the present study, lactic acid bacteria and propionic acid bacteria have been identified at the genus level with the use of artificial neural networks (ANNs) and Fourier transform infrared spectroscopy (FTIR). Bacterial strains of the genera *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Streptococcus* and *Propionibacterium* were analyzed since they deliver health benefits and are routinely used in the food processing industry. The correctness of bacterial identification by ANNs and FTIR was evaluated at two stages. At first stage, ANNs were tested based on the spectra of 66 reference bacterial strains. At second stage, the evaluation involved 286 spectra of bacterial strains isolated from food products, deposited in our laboratory collection, and identified by genus-specific PCR. ANNs were developed based on the spectra and their first derivatives. The most satisfactory results were reported for the probabilistic neural network, which was built using a combination of W5W4W3 spectral ranges. This network correctly identified the genus of 95 % of the lactic acid bacteria and propionic acid bacteria strains analyzed.

*Key words*: lactic acid bacteria, propionic acid bacteria, FTIR spectroscopy, artificial neural networks

## Introduction

The correct identification of bacteria plays a vital role in microbial diagnostics. New methods supporting the identification of microorganisms have been developed. At present, the main focus is on the simplicity and time efficiency of the proposed analytical procedures while maintaining high specificity and accuracy of the results. Fourier transform infrared spectroscopy (FTIR) combined with artificial neural networks (ANNs) meets the above requirements. FTIR spectra are strain specific, and they reveal the characteristic features of all cellular components, such as fatty acids, membrane proteins, intracellular proteins, polysaccharides and nucleic acids. According to Naumann *et al.* (*1*), for identification purposes, five spectral regions in IR spectra can be distinguished. They are called spectral 'windows': W1 (3000–2800 $cm^{-1}$) is the fatty acid region; W2 (1700–1500 $cm^{-1}$) contains the amide I and II bands of proteins and peptides; W3 (1500–1200 $cm^{-1}$) is a mixed region of fatty acid bending vibrations, proteins, and phosphate-carrying compounds; W4 (1200–900 $cm^{-1}$) contains absorption bands of the carbohydrates in microbial cell walls; W5 (900–700 $cm^{-1}$) is the 'fingerprint region' that contains weak but very unique absorbancies that are characteristic of specific bacteria. The differences between microbial spectra are difficult to observe, which is why they have to be analyzed with the use of multivariate statistical methods, among which the most often applied are cluster analysis, discriminant analysis, discriminant function analysis or canonical variate analysis (*2*,*3*). In addition to traditional methods, microbial diversity is also determined with the involvement of ANNs (*4*–*9*). ANNs are mathematical models which process data on a pseudo–parallel

*Corresponding author; Phone: ++48 89 523 3687; Fax: ++48 89 523 4945; E-mail: bartlomiej.dziuba@uwm.edu.pl

basis and consist of an interconnected group of artificial neurons that imitate the activity of biological brain structures (*10*). Their development is the result of an interdisciplinary merger of conventional sciences, including biology, physics and mathematics. Neural networks differ from data processing algorithms in that they are capable of generalizing the learned knowledge to the data that have not been previously known or presented in the course of system training.

In our previous studies (*11–15*), we presented a methodology for measuring FTIR spectra and a strategy for spectral analyses used to identify lactic acid bacteria at the genus or species level. To date, the use of ANNs and FTIR spectroscopy as the most advanced chemometric methods for identifying bacteria of the genus *Propionibacterium* has not been discussed in the existing body of literature.

Propionic acid bacteria are used in the production of cheese, silage, fermented vegetable products, skimmed milk for calves, probiotics as well as protein and vitamin supplements used in animal nutrition (*16*,*17*). *Propionibacteria* are found in commercially available products that inhibit the growth of pathogens in dairy foods (*e.g.* MicroGARD™, Danisco, Copenhagen, Denmark) and animal silage (*e.g.* AIV Bioprofit™, Kemira GrowHow, Helsinki, Finland) (*18*). The application of *Propionibacteria* in the food industry is due to their specific properties: they support lactose and lactate utilization; are characterized by the activity of cell membrane proteinase and intracellular peptidases; synthesize compounds that give the product its characteristic aroma, flavour and texture; synthesize antibacterial and antifungal compounds as well as $CO_2$; and produce vitamin B12 and folic acid (*19*). Selected *Propionibacterium* species were given a GRAS status, therefore, the produced vitamins and bacterial cells may be safely added to foods.

Bacteria of the genus *Propionibacterium* are employed mainly in the production of hard rennet cheese (Swiss cheese: Emmental, Gruyère, Appenzeller, Raclettes; Dutch cheese: Leerdammer, Maasdammer; Norwegian cheese: Jarlsberg; French cheese: Comté) (*20*). They proliferate slowly during cheese fermentation and ripening. The acids that are relatively quickly produced by lactic acid bacteria reduce the cheese pH, thus inhibiting the growth of *Propionibacterium*. They begin to proliferate only when lactates are produced and acidity decreases. They convert lactates to propionic acid and acetic acid. The above compounds give cheese its characteristic aroma, and they act as natural food preservatives (*21*).

In the present study, lactic acid bacteria and propionic acid bacteria have been identified at the genus level with the use of ANNs and FTIR. Analysis was focused on bacterial strains of the genera *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Streptococcus* and *Propionibacterium* that deliver health benefits and are popularly used in the food processing industry. This study also made an attempt to expand the library of FTIR spectra of microorganisms (*11*).

## Materials and Methods

### *Bacterial strains*

Sixty-six reference strains of lactic acid bacteria and propionic acid bacteria from international collections, including American Type Culture Collection (ATCC), Belgian Coordinated Collections of Microorganisms/LMG Bacteria Collection (LMG), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ) (Table 1), and 306 strains isolated from food products and supplied by the Department of Industrial and Food Microbiology at the University of Warmia and Mazury in Olsztyn, Poland, were included in this study. The taxonomic status of food isolates identified previously by classical phenotypic methods was: 152 *Lactobacillus*, 30 *Propionibacterium*, 93 *Lactococcus*, 21 *Leuconostoc* and 10 *Streptococcus thermophilus* strains.

Table 1. List of reference strains

| Reference strains | |
|---|---|
| *Lactobacillus acidophilus* | DSMZ 20079 *(T)*, DSMZ 9126, LMG 9433 |
| *Lb. amylolyticus* | DSMZ 11664 *(T)* |
| *Lb. amylovorus* | DSMZ 20531 *(T)* |
| *Lb. brevis* | DSMZ 20054 *(T)*, DSMZ 1267, LMG 6906 |
| *Lb. casei* | DSMZ 20011 *(T)*, LMG 6904 |
| *Lb. delbrueckii* ssp. *bulgaricus* | DSMZ 20081*(T)*, DSMZ 20080, LMG 7942, DSMZ 20074 *(T)*, ATCC 9649, DSMZ 20072 *(T)*, DSMZ 20355, ATCC 7830 |
| *Lb. fermentum* | DSMZ 20052 *(T)*, DSMZ 20391, LMG 6902 |
| *Lb. helveticus* | DSMZ 20075 *(T)*, LMG 6413, ROS |
| *Lb. hilgardii* | DSMZ 20176 *(T)*, DSMZ 20051, |
| *Lb. jensenii* | DSMZ 20557 *(T)* |
| *Lb. paracasei* ssp. *paracasei* | DSMZ 5622 *(T)*, DSMZ 2649 |
| *Lb. paraplantarum* | DSMZ 10667 *(T)*, DSMZ 10641 |
| *Lb. plantarum* ssp. *plantarum* | DSMZ 20174 *(T)* |
| *Lb. rhamnosus* | ATCC 7469, VKM B-574 |
| *Lb. zeae* | DSMZ 20178 *(T)* |
| *Lb. pentosus* | DSMZ 20314 *(T)*, DSMZ 16366 |
| *Lactococcus lactis* ssp. *cremoris* | DSMZ 20069 *(T)*, DSMZ 4367 |
| *Lc. lactis* ssp. *lactis* | DSMZ 20481*(T)*, DSMZ 4366 |
| *Lc. raffinolactis* | DSMZ 20443 *(T)* |
| *Lc. plantarum* | DSMZ 20686 *(T)*, DSMZ 20687 |
| *Leuconostoc lactis* | DSMZ 20202 *(T)*, DSMZ 8581 |
| *Ln. pseudomesenteroides* | DSMZ 20193 *(T)*, DSMZ 5624 |
| *Ln. mesenteroides* ssp. *mesenteroides* | DSMZ 20343 *(T)*, DSMZ 20240 |
| *Ln. mesenteroides* ssp. *dextranicum* | DSMZ 20484 *(T)*, DSMZ 20071 |
| *Ln. mesenteroides* ssp. *cremoris* | DSMZ 20346 *(T)*, DSMZ 20200 |
| *Propionibacterium acidipropionici* | DSMZ 4900 *(T)*, DSMZ 20272 |
| *P. freudenreichii* ssp. *freudenreichii* | DSMZ 20271*(T)* |
| *P. freudenreichii* ssp. *shermanii* | DSMZ 4902 *(T)*, DSMZ 20270 |
| *P. jensenii* | DSMZ 20535 *(T)*, DSMZ 20274 |
| *P. thoenii* | DSMZ 20276 *(T)*, DSMZ 20275 |
| *Streptococcus thermophilus* | NCFB, SY102, CH101 |

### *Culture cultivation*

Strains were cultivated on solid media for (48±2) h under optimum conditions: mesophilic and thermophilic lactobacilli on MRS agar (Merck, Darmstadt, Germa-

ny) at 30 and 37 °C, respectively; *Leuconostoc* strains on MRS agar (Merck) at 25 °C; *Streptococcus thermophilus* strains on M17 agar (Merck) at 42 °C; *Lactococcus* on M17 agar (Merck) at 30 °C; *Propionibacteria* on agar medium with calcium lactate (Sigma-Aldrich, St. Louis, MO, USA) as the only source of carbon at 30 °C.

### *DNA isolation and PCR amplification*

Loopful amount of the culture was taken from the agar plate, suspended in 100 μL of Tris-HCl (10 mM, pH=8.5) buffer and incubated at 37 °C for about 1 h in the presence of lysozyme added to a final concentration of 2 mg/mL. Further steps were done according to the manufacturer's instructions for Genomic Mini kit (A&A Biotechnology, Gdynia, Poland) for genomic DNA isolation. Generally, the method is based on the DNA ability to bind to silica gels in the presence of high concentrations of chaotropic salts. PCR methodology relies on primers and procedures developed by Dubernet *et al*. (*22*), Macián *et al*. (*23*) and Rossi *et al*. (*24*) for genus identification of *Lactobacillus*, *Leuconostoc* and *Propionibacterium* strains, respectively. *Lactococcus* strains were identified with species-specific primers as described by Pu *et al*. (*25*). *Streptococcus thermophilus* strains had previously been identified in the course of EU INCO–Copernicus IC15--CT98-0905 project. Primers and the profiles of the PCR amplifications, as well as the expected product size are presented in Table 2. Amplification was carried out in a thermal cycler Bio-Rad MJ Mini (Bio-Rad Laboratories, Hercules, CA, USA). The reaction mixture (20 μL) contained 10 ng of each primer (Oligo, Warsaw, Poland), 0.2 mM of each dNTP (Fermentas International, Vilnius, Lithuania), 1×PCR buffer with 2.5 mM $MgCl_2$, depending on the experiment 20–40 ng of bacterial DNA and 1–2 U of Taq DNA Polymerase (Fermentas International).

### *Spectroscopic measurements*

Bacterial strain samples were subjected to spectroscopic measurements by the transmission method within the wavelength range of 4000 to 500 $cm^{-1}$ using a FTIR spectrophotometer (Spectrum One, PerkinElmer, Inc, Waltham, MA, USA) equipped with a beam splitter (KBr) and a deuterated triglycine sulphate (DTGS) detector. Every sample was scanned 64 times at the resolution of 4 $cm^{-1}$ and scanning speed of 0.5 cm/s. The background spectrum was measured with an empty ZnSe window (PerkinElmer, Inc). The applied spectral operations and statistical analyses of spectral data have been described in detail in our previous work (*13*).

### *Development of artificial neural networks*

Two types of ANNs were used: probabilistic neural networks (PNNs) (*26*) and multilayer perceptrons (MLPs) (*27*). The first layer of the network was an input layer, to which the absorbance values at a given wavelength were introduced. The number of neurons in the input layer was equal to the number of data points of the spectrum or selected wave numbers. The spectral data were divided into three sets (learning, validation and testing) to include the spectra of all reference strains analyzed. The hidden layer consisted of neurons whose number was equal to the number of learning examples in the case of PNNs, and defined experimentally in that of MLPs. The output layer consisted of five neurons, each of which represented a given genus.

The multilayer perceptrons used in the study consisted of three to four layers. Back–propagation and conjugate gradient descent algorithms were used in the learning process, which was conducted until the root mean square (RMS) error between the actual and desired outputs was lower than 0.05.

Probabilistic neural networks are a type of the so--called Bayesian networks that use kernel–based approximation to form an estimate of the probability density functions of classes in a classification problem. The PNN used in this study consisted of three layers (input, radial and linear) of classification neurons. The only control factor that needs to be selected for probabilistic neural

Table 2. List of specific primers for PCR identification of isolates

| Primer | Sequence 5'–3' | Specifity | PCR reaction profile | PCR product/bp |
|---|---|---|---|---|
| LbLMA1<br>R16-1 | CTCAAAACTAAACAAAGTTTC<br>CTTGTACACACCGCCCGTCA | *Lactobacillus* sp. | 95 °C – 5'<br>95 °C – 30"<br>30× 55 °C – 30"<br>72 °C – 30"<br>72 °C – 7' | ~250 |
| LeucA<br>LeucS | CACTTTGTCTCCGAAGAG<br>AAGCACTGTTGTATGGGA | *Leuconostoc* sp. | 94 °C – 5'<br>94 °C – 30"<br>35× 56 °C – 45"<br>72 °C – 45"<br>72 °C – 5' | 613 |
| PB1<br>PB2 | AGTGGCGAAGGCGGTTCTCTGGA<br>TGGGGTCGAGTTGCAGACCCCAAT | *Propionibacterium* sp. | 94 °C – 4'<br>94 °C – 30"<br>40× 68 °C – 15"<br>72 °C – 1'<br>72 °C – 5' | 610 |
| 1RL<br>LacreR<br>LgR<br>PiplraR | TTTGAGAGTTTGATCCTGG<br>GGGATCATCTTTGAGTGAT<br>AAGTAATTTTCCACTCTACTT<br>CGTCACTGAGGGCTGGAT | *Lactococcus lactis*<br>*Lc. garvieae*<br>*Lc. plantarum,*<br>*Lc. piscium,*<br>*Lc. raffinolactis* | 94 °C – 4'<br>94 °C – 30"<br>35× 45 °C – 30"<br>72 °C – 2'<br>72 °C – 2' | 238<br>482<br>860 |

network training is the smoothing factor (*i.e.* the radial deviation of the Gaussian function), which was set between 0.3 and 1.0.

All artificial neural networks were developed in the STATISTICA Neural Network v. 4.01 application (*10*). The correctness of bacterial identification was verified using molecular biology methods (PCR).

## Results and Discussion

The reproducibility and comparability of results in the long–term perspective, which are important considerations in the use of FTIR spectroscopy for the identification and classification of lactic acid bacteria and propionic acid bacteria, were evaluated in respect to microbial variability, experimental artefacts and equipment defects. A quantitative measure had to be applied to determine the quality of bacterial FTIR spectra. The differentiation index *D*, proposed by Naumann *et al.* (*1*) and calculated based on Pearson's correlation coefficients, was used in this study. The cited authors (*1*) have demonstrated that spectra cannot be differentiated if the value of index *D*, calculated based on the first derivatives of FTIR spectra in the range of 4000–500 $cm^{-1}$, is less than 10.

To evaluate the effect of bacterial culture conditions and sample preparation methods on the quality of bacterial FTIR spectra, the reproducibility of measurements was analyzed at two levels (Table 3). Reproducibility level 1 (RL1) indicates the reproducibility of spectral measurements for the samples prepared from the same bacterial cell suspension, whereas RL2 describes the reproducibility of spectral measurements for samples prepared from independent cultures of the same bacterial strain cultivated under the same conditions. RL2 had practical implications for the performed measurements, although in many cases it was significantly lower than RL1. Based on the analysis of *D* values, the FTIR spectra of the analyzed bacteria were selected and used to supplement the library of microbial spectra. The first derivatives of FTIR spectra were most effective in differentiating lactic acid bacteria and propionic acid bacteria at the genus level. The most satisfactory results were observed in the dactyloscopic region (W5) (900–700 $cm^{-1}$), the polysaccharide region (W4) (1200–900 $cm^{-1}$) and the combination region (W3) (1500–1200 $cm^{-1}$). The results were processed by a multivariate statistical analysis to validate the correctness of conclusions derived from a comparison of bacterial spectra based on the values of *D*. This analysis also revealed combinations of spectral ranges that supported the generation of the most accurate results. The analyzed lactic acid bacteria and propionic acid bacteria were most effectively differentiated and identified at the genus level in a combination of W5W4W3 spectral ranges. Our observations are largely consistent with the results reported by other authors in respect of lactic acid bacteria (*2,4,28–30*). Data regarding analyses of FTIR spectra of propionic acid bacteria are not available.

The correctness of bacterial identification by ANNs and FTIR was evaluated at two stages, according to the strategy described in our previous works (*12,14,15*). At the first stage, artificial neural networks were tested based on the spectra of 66 reference bacterial strains (Table 1). At the second stage, the evaluation involved 286 spectra of bacterial strains isolated from food products and deposited in the collection of the Department of Industrial and Food Microbiology at the University of Warmia and Mazury in Olsztyn, Poland, identified by PCR screening with the use of genus- or species-specific primers. Ten strains of *Streptococcus thermophilus* were excluded from PCR screening as they had been identified previously. Examples of agarose gel electrophoresis of PCR amplification products specific for *Propionibacterium*, *Lactobacillus*, *Leuconostoc* and *Lactococcus lactis* strains are shown in Fig. 1. An analysis of PCR results with the use of primers specific for the examined genera of lactic acid bac-

Table 3. Reproducibility of FTIR measurements for the spectra of *Lactococcus*, *Leuconostoc*, *Lactobacillus*, *Streptococcus* and *Propionibacterium* strains

| Spectrum range/$cm^{-1}$ | Reproducibility level | *Lactococcus* | | *Leuconostoc* | | *Lactobacillus* | | *Streptococcus* | | *Propionibacterium* | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | FTIR spectrum | I derivative | FTIR spectrum | I derivative | FTIR spectrum | I derivative | FTIR spectrum | I derivative | FTIR spectrum | I derivative |
| 4000–500 | RL1 | 0.7 | 2.8 | 0.5 | 3.3 | 1.4 | 4.5 | 1.7 | 5.2 | 1.3 | 4.2 |
| | RL2 | 3.4 | 6.8 | 2.2 | 7.3 | 7.5 | 9.4 | 6.4 | 8.8 | 4.6 | 8.4 |
| 3100–2800 | RL1 | 0.3 | 1.0 | 0.2 | 0.4 | 0.3 | 1.2 | 0.4 | 0.8 | 1.3 | 3.1 |
| | RL2 | 1.6 | 3.5 | 1.1 | 2.8 | 1.4 | 3.7 | 1.4 | 3.5 | 2.6 | 5.1 |
| 1800–1500 | RL1 | 0.3 | 2.7 | 0.6 | 1.4 | 1.2 | 3.3 | 1.2 | 2.6 | 1.5 | 4.2 |
| | RL2 | 1.9 | 7.5 | 2.2 | 10.7 | 3.8 | 11.2 | 3.4 | 10.6 | 4.7 | 12.2 |
| 1500–1200 | RL1 | 2.2 | 4.9 | 0.8 | 2.4 | 1.8 | 3.3 | 1.6 | 3.3 | 2.9 | 3.8 |
| | RL2 | 5.6 | 10.4 | 3.5 | 8.8 | 4.0 | 9.6 | 5.8 | 10.7 | 6.4 | 6.8 |
| 1200–900 | RL1 | 0.5 | 1.0 | 0.1 | 0.7 | 0.8 | 2.9 | 0.6 | 1.6 | 0.4 | 0.7 |
| | RL2 | 2.5 | 6.6 | 2.6 | 6.1 | 2.9 | 7.5 | 1.9 | 7.8 | 2.4 | 9.8 |
| 900–700 | RL1 | 2.1 | 4.9 | 0.9 | 3.6 | 8.8 | 8.4 | 8.1 | 6.2 | 7.9 | 5.1 |
| | RL2 | 7.9 | 11.4 | 8.7 | 10.2 | 9.6 | 12.1 | 11.3 | 12.8 | 11.1 | 12.3 |

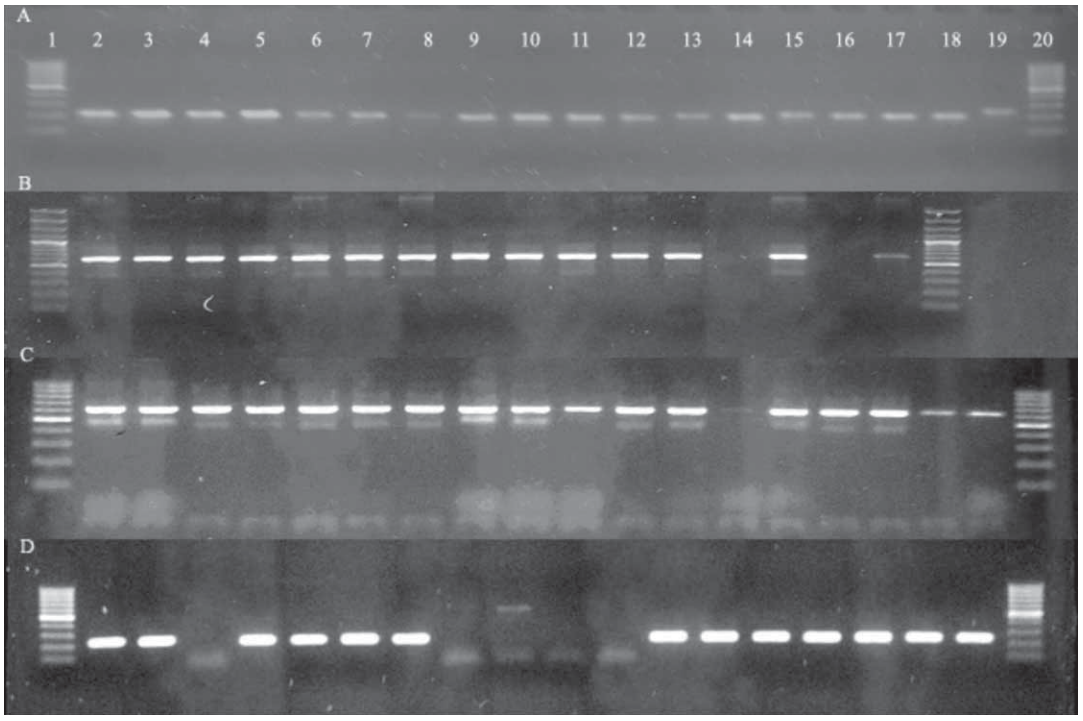


**Fig. 1.** Agarose gel electrophoresis of PCR products obtained with primers: (A) LbLMA1-rev and R16-1 specific for *Lactobacillus* spp., product ~250 bp. Lanes: 2 – strain 1G, 3 – strain T44, 4 – strain 22M, 5 – strain 25G, 6 – strain T53/1, 7 – strain T106/3, 8 – strain 701, 9 – strain T217, 10 – strain T106/2, 11 – strain T106/1, 12 – strain T116/2, 13 – strain T116/1, 14 – strain T48/2, 15 – strain T45, 16 – strain 134, 17 – strain T8/4, 18 – strain T135/2, 19 – strain T135/4; (B) PB1 and PB2 specific for *Propionibacterium* spp., product 610 bp. Lanes: 2 – strain P112/2, 3 – strain P109/1, 4 – strain 113/1, 5 – strain P2/2, 6 – strain P110C, 7 – strain P124, 8 – strain P109B, 9 – strain P84, 10 – strain P116, 11 – strain P83, 12 – strain P123/2, 13 – strain P109B1, 14 – strain C08, 15 – strain P109C, 16 – strain P127/5, 17 – strain 119; (C) LeucA and LeucS specific for *Leuconostoc* spp., product ~613 bp. Lanes: 2 – strain Iz33, 3 – strain Iz31, 4 – strain Iz39, 5 – strain Iz13, 6 – strain Iz14, 7 – strain Iz1, 8 – strain Iz2, 9 – strain Iz3, 10 – strain T48, 11 – strain T146, 12 – strain 65/2, 13 – strain 68/1, 14 – strain 62, 15 – strain 3/1, 16 – strain L4, 17 – strain L7, 18 – strain L3, 19 – strain T46; (D) 1RL and LacreR specific for *Lactococcus lactis*, product 238 bp. Lanes: 2 – strain C35, 3 – strain 4, 4 – strain 55/2.2L4, 5 – strain 55/2.1, 6 – strain 5, 7 – strain 53/2, 8 – strain 672, 9 – strain 56/1, 10 – strain 2, 11 – strain T68/2, 12 – strain 44/1, 13 – strain 63, 14 – strain 51/1, 15 – strain 53/1, 16 – strain 78, 17 – strain 3, 18 – strain 9, 19 – strain 1. Lanes 1 and 20, and 18 for B – mass marker O'GeneRuler 100bp DNA Ladder

teria and propionic acid bacteria revealed that in the group of strains isolated from food products and initially classified by classical phenotypic methods as *Lactobacillus* (152 strains), *Propionibacterium* (30 strains), *Lactococcus* (93 strains) and *Leuconostoc* (21 strains), the genus identity of 136, 28, 87 and 18 strains, respectively, has been validated. Six strains stated as *Lactococcus* were then identified as *Leuconostoc*, one strain of *Leuconostoc* as *Lactococcus*. Strains not identified as *Lactobacillus* (16 strains), *Propionibacterium* (2 strains) and *Leuconostoc* (2 strains) were removed from that research and not further examined.

The FTIR spectra of bacteria comprise hundreds or even thousands of overlapping absorption bands that are impossible to separate. Their analysis requires image recognition methods which examine spectra as fingerprints (dactyloscopic images). ANNs are among the most advanced analytical tools. In this study, ANNs were developed with the use of spectra and their first derivatives. The search for input variables and the training process led to the selection of the best neural networks which are presented in Table 4. Those networks correctly identified all of the reference strains (quality of test 1 data, first verification stage) and 89 to 95 % of the remaining

Table 4. Set of the best artificial neural networks used for genus identification of selected lactic acid bacteria and *Propionibacterium* strains

| No. | FTIR spectrum range | ANN type | ANN architecture | RMS | | | Test 2 quality | No. of epochs |
|---|---|---|---|---|---|---|---|---|
| | | | | Learning | Validation | Test 1 | % | |
| 1 | 4000–500 $cm^{-1}$ | MLP | 1049:36:5 | 0.006 | 0.01 | 0.08 | 0.92 | $6{\cdot}10^3$ |
| 2 | W5W4W3W2 | MLP | 389:32:5 | 0.00 | 0.01 | 0.02 | 0.89 | $3{\cdot}10^2$ |
| 3 | AL.GEN | MLP | 1001:36:5 | 0.00 | 0.02 | 0.02 | 0.94 | $8{\cdot}10^3$ |
| 4 | W5W4W2 | PNN | 800:740:5 | 0.00 | 0.006 | 0.002 | 0.91 | – |
| 5 | W5W4W3 | PNN | 800:740:5 | 0.00 | 0.003 | 0.003 | 0.95 | – |

Table 5. Summary of genus typing by PNN (800:740:5)

| Genus | No. of strains | Typing by PNN | | |
|---|---|---|---|---|
| | | Correct | No answer | Wrong |
| *Lactobacillus* spp. | 136 | 127 | 2 | 7 |
| *Lactococcus* spp. | 88 | 86 | 1 | 1 |
| *Leuconostoc* spp. | 24 | 23 | 1 | 0 |
| *Streptococcus* spp. | 10 | 10 | 0 | 0 |
| *Propionibacterium* spp. | 28 | 24 | 2 | 2 |

Table 6. Problematic or false typing of strain genus by PNN (800:740:5)

| Strain | PCR typing | ANN typing | Unit error | Activation value on output neurons | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | *P.* | *Lb.* | *Ln.* | *Lc.* | *St.* |
| P109/1 | *Propionibacterium* sp. | ? | 0.26 | 0.59 | 0.00 | 0.00 | 0.41 | 0.00 |
| P84 | *Propionibacterium* sp. | ? | 0.24 | 0.62 | 0.00 | 0.00 | 0.38 | 0.00 |
| P78 | *Propionibacterium* sp. | *Ln.* | 0.63 | 0.00 | 0.02 | 0.98 | 0.00 | 0.00 |
| P20C | *Propionibacterium* sp. | *Ln.* | 0.63 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 |
| 9 | *Lactococcus* sp. | ? | 0.41 | 0.00 | 0.00 | 0.00 | 0.35 | 0.65 |
| 71/1 | *Lactococcus* sp. | *Ln.* | 0.63 | 0.00 | 0.00 | 0.70 | 0.30 | 0.00 |
| 3/1 | *Leuconostoc* sp. | ? | 0.41 | 0.00 | 0.00 | 0.35 | 0.00 | 0.65 |
| 22M | *Lactobacillus* sp. | ? | 0.35 | 0.00 | 0.46 | 0.54 | 0.00 | 0.00 |
| T45 | *Lactobacillus* sp. | *Lc.* | 0.45 | 0.00 | 0.45 | 0.00 | 0.55 | 0.00 |
| T103 | *Lactobacillus* sp. | ? | 0.20 | 0.00 | 0.31 | 0.69 | 0.00 | 0.00 |
| Ros | *Lactobacillus* sp. | *Ln.* | 0.25 | 0.00 | 0.50 | 0.50 | 0.00 | 0.00 |
| L200 | *Lactobacillus* sp. | *Ln.* | 0.60 | 0.00 | 0.12 | 0.88 | 0.00 | 0.00 |
| L206 | *Lactobacillus* sp. | *Ln.* | 0.48 | 0.00 | 0.16 | 0.84 | 0.00 | 0.00 |
| 4M | *Lactobacillus* sp. | *Ln.* | 0.26 | 0.00 | 0.22 | 0.78 | 0.00 | 0.00 |
| LPC57 | *Lactobacillus* sp. | *Ln.* | 0.29 | 0.00 | 0.29 | 0.71 | 0.00 | 0.00 |
| AC/17 | *Lactobacillus* sp. | *Ln.* | 0.26 | 0.00 | 0.19 | 0.81 | 0.00 | 0.00 |

*P.=Propionibacterium, Lb.=Lactobacillus, Ln.=Leuconostoc, Lc.=Lactococcus, St.=Streptococcus*

286 strains (quality of test 2 data). The best results were reported for the probabilistic network (Table 4, No. 5) developed based on a combination of W5W4W3 spectral ranges. The network correctly identified 95 % lactic acid bacteria and propionic acid bacteria at the genus level. In the group of 286 tested strains, the network unambiguously determined the genus of 270 strains (Table 5). Problematic or falsely typed strains are presented in Table 6. Two *Propionibacterium* strains – P109/1 and P84, one *Lactococcus* strain – 9, one *Leuconostoc* strain – 3/1, and two *Lactobacillus* strains – 22M and T103, were not unambiguously identified. As regards the winning neuron, *Propionibacterium* strains were correctly classified, whereas the strains of the genera *Lactococcus*, *Leuconostoc* and *Lactobacillus* were incorrectly identified. The following strains were incorrectly classified: P78 and P20C of the genus *Propionibacterium* – as belonging to *Leuconostoc*, 71/1 of the genus *Lactococcus* as belonging to *Leuconostoc*, T45 of the genus *Lactobacillus* as belonging to *Lactococcus*, and strains Ros, L200, L206, 4M, LPC57 and AC/17 were recognized as belonging to *Leuconostoc.*

The present method may be deployed in analytical laboratories for identifying lactic and propionic acid bacteria, for monitoring the purity of cultures in strain collections and for fast screening of selected bacterial groups. The proposed technique is also suitable for monitoring the quality of products and raw materials in the food processing industry (*31,32*). FTIR delivers a variety of advantages, including simple technology, low cost, high specificity and a wide range of industrial applications. The method could be expanded to include successive neural networks for identifying pathogenic bacteria (*33*).

Our study will support the development of a database panel (FTIR spectra of the analyzed bacteria) and an analytical panel for identifying bacteria with the use of ANNs. The resulting BACTI-FTIR database will be available at *http://www.uwm.edu.pl/mikrobiologia* to all persons interested in the identification of bacteria based on FTIR spectra, with the use of ANNs.

## Conclusions

The use of artificial neural networks as the most advanced chemometric method for analzying FTIR spectra supported the correct identification of 95 % *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Streptococcus* and *Propionibacterium* bacteria at the genus level under the conditions applied in this study. RL2, which describes the reproducibility of spectral measurements for samples prepared from independent cultures of the same bacterial strain, has practical implications for the measurement of FTIR spectra of the studied bacteria. For identification purposes the most suitable were the first derivatives of FTIR spectra within the polysaccharide region (1200–900 $cm^{-1}$), the dactyloscopic region (900–700 $cm^{-1}$) and the combined region (1500–1200 $cm^{-1}$). The results of this study will support the development of the BACTI–FTIR database, which will be regularly expanded and updated to offer a simple and practical tool for identifying microorganisms.

*Acknowledgements*

This work was supported by the Ministry of Science and Higher Education, Poland, within project no. N312 009 31/1065.

## References

1. D. Naumann, H. Labischinski, P. Giesbrecht: The Characterization of Microorganisms by Fourier-Transform Infrared Spectroscopy (FT-IR). In: *Modern Techniques for Rapid Microbiological Analysis*, W.H. Nelson (Ed.), VCH Publishers, New York, NY, USA (1991) pp. 43–96.
2. J. Samelis, A. Bleicher, C. Delbès-Paus, A. Kakouri, K. Neuhaus, M.C. Montel, FTIR-based polyphasic identification of lactic acid bacteria isolated from traditional Greek Graviera cheese, *Food Microbiol. 28* (2011) 76–83.
3. C. Amiel, L. Mariey, M.C. Curk–Daubié, P. Pichon, J. Travert, Potentiality of fourier transform infrared spectroscopy (FTIR) for discrimination and identification of dairy lactic acid bacteria, *Lait, 80* (2000) 445–459.
4. C. Amiel, L. Mariey, C. Denis, P. Pichon, J. Travert, FTIR spectroscopy and taxonomic purpose: Contribution to the classification of lactic acid bacteria, *Lait, 81* (2001) 249–255.
5. R. Goodacre, E.M. Timmins, R. Burton, N. Kaderbhai, A.M. Woodward, D.B. Kell P.J. Rooney, Rapid identification of urinary tract infection bacteria using hyperspectral whole-organism fingerprinting and artificial neural networks, *Microbiology, 144* (1998) 1157–1170.
6. C. Kirschner, K. Maquelin, P. Pina, N.A. Ngo-Thi, L.P. Choo-Smith, G.D. Sockalingum, Classification and identification of enterococci: A comparative phenotypic, genotypic, and vibrational spectroscopic study, *J. Clin. Microbiol. 39* (2001) 1763–1770.
7. K. Tintelnot, G. Haase, M. Seibold, F. Bergmann, M. Staemmler, T. Franz, D. Naumann, Evaluation of phenotypic markers for selection and identification of *Candida dubliniensis*, *J. Clin. Microbiol. 38* (2000) 1599–1608.
8. T. Udelhoven, D. Naumann, J. Schmitt, Development of hierarchical classification systems with artificial neural networks and FT-IR spectra for the identification of bacteria, *Appl. Spectrosc. 54* (2000) 1471–1479.
9. M. Wenning, N.R. Büchl, S. Scherer, Species and strain identification of lactic acid bacteria using FTIR spectroscopy and artificial neural networks, *J. Biophotonics, 3* (2010) 493–505.
10. R. Tadeusiwicz, P. Lula, STATISTICA Neural Network v. 4.01: Course Materials, StatSoft Inc, Warsaw, Poland (2002).
11. B. Dziuba, A. Babuchowski, M. Niklewicz, B. Brzozowski, FTIR spectral characteristics of lactic acid bacteria: A spectral library, *Milchwissenschaft, 61* (2006) 146–149.
12. B. Dziuba, A. Babuchowski, M. Dziuba, D. Nałęcz, Identification of lactic acid bacteria using FTIR spectroscopy and artificial networks, *Milchwissenschaft, 62* (2007) 28–32.
13. B. Dziuba, A. Babuchowski, D. Nałęcz, M. Niklewicz, Identification of lactic acid bacteria using FTIR spectroscopy and cluster analysis, *Int. Dairy J. 17* (2007) 183–189.
14. B. Dziuba, Identification of *Lactobacillus* strains at the species level using FTIR spectroscopy and artificial neural networks, *Pol. J. Food. Nutr. Sci. 57* (2007) 301–306.
15. B. Dziuba, Identification of selected *Leuconostoc* species with the use of FTIR spectroscopy and artificial neural networks, *Acta Sci. Pol. Technol. Aliment. 10* (2011) 275–285.
16. I. Filya, E. Sucu, A. Karabulut, The effect of *Propionibacterium acidipropionici*, with or without *Lactobacillus plantarum*, on the fermentation and aerobic stability of wheat, sorghum and maize silages, *J. Appl. Microbiol. 97* (2004) 818–826.
17. R.J Merry, D.R. Davies, Propionibacteria and their role in the biological control of aerobic spoilage in silage, *Lait, 79* (1999) 149–164.
18. H. Holo, T. Faye, D.A. Brede, T. Nilsen, I. Odegard, T. Langsrud *et al.*, Bacteriocins of propionic acid bacteria, *Lait, 82* (2002) 59–68.
19. J. Hugenholtz, J. Hunik, H. Santos, E. Smid, Nutraceutical production by propionibacteria, *Lait, 82* (2002) 103–112.
20. K. Hojo, R. Watanabe, T. Mori, N. Taketomo, Quantitative measurement of tetrahydromenaquinone-9 in cheese fermented by propionibacteria, *J. Dairy Sci. 90* (2007) 4078–4083.
21. D.F.M. Daly, P.L.H. McSweeney, J.J. Sheehan, Split defect and secondary fermentation in Swiss-type cheeses – A review, *Dairy Sci. Technol. 90* (2010) 3–26.
22. S. Dubernet, N. Desmasures, M. Guéguen, A PCR-based method for identification of lactobacilli at genus level, *FEMS Microbiol. Lett. 214* (2002) 271–275.
23. M.C. Macián, E. Chenoll, R. Aznar, Simultaneous detection of *Carnobacterium* and *Leuconostoc* in meat products by multiplex PCR, *J. Appl. Microbiol. 97* (2004) 384–394.
24. F. Rossi, S. Torriani, F. Dellaglio, Genus- and species-specific PCR-based detection of dairy propionibacteria in environmental samples by using primers targeted to the genes encoding 16S rRNA, *Appl. Environ. Microbiol. 65* (1999) 4241–4244.
25. Z.Y. Pu, M. Dobos, G.K.Y. Limsowtin, I.B. Powell, Integrated polymerase chain reaction-based procedures for the detection and identification of species and subspecies of the Gram-positive bacterial genus *Lactococcus*, *J. Appl. Microbiol. 93* (2002) 353–361.
26. D.F. Specht, Probabilistic neural networks, *Neural Networks, 3* (1990) 109–118.
27. C. Bishop: *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, UK (1995).
28. L. Mariey, J.P. Signolle, C. Amiel, J. Travert, Discrimination, classification, identification of microorganisms using FTIR spectroscopy and chemometrics, *Vib. Spectrosc. 26* (2001) 151–159.
29. K. Maquelin, C. Kirschner, L.P. Choo-Smith, N.A. Ngo-Thi, T. van Vreeswijk, M. Stämmler *et al.*, Prospective study of the performance of vibrational spectroscopies for rapid identification of bacterial and fungal pathogens recovered from blood cultures, *J. Clin. Microbiol. 41* (2003) 324–329.
30. B. Weinrichter, W. Luginnbühl, H. Rohm, J. Jimeno, Differentiation of facultatively heterofermentative lactobacilli from plants, milk, and hard type cheeses by SDS-PAGE, RAPD, FTIR, energy source utilization and autolysis type, *Lebensm. Wiss. Technol. 34* (2001) 556–566.
31. E. Garcia-Caurel, J. Nguyen, L. Schwartz, B. Drévillon, Application of FTIR ellipsometry to detect and classify microorganisms, *Thin Solid Films, 455–456* (2004) 722–725.
32. S. Goerges, J. Mounier, M.C. Rea, R. Gelsomino, V. Heise, R. Beduhn *et al.*, Commercial ripening starter microorganisms inoculated into cheese milk do not successfully establish themselves in the resident microbial ripening consortia of a South German red smear cheese, *Appl. Environ. Microbiol. 74* (2008) 2210–2217.
33. C.A. Rebuffo, J. Schmitt, M. Wenning, F. von Stetten, S. Scherer, Reliable and rapid identification of *Listeria monocytogenes* and *Listeria* species by artificial neural network–based Fourier transform infrared spectroscopy, *Appl. Environ. Microbiol. 72* (2006) 994–1000.