Application of MALDI-TOF MS for the subtyping of *Arcobacter butzleri* strains and comparison with their MLST and PFGE types

Application of MALDI-TOF MS for the subtyping of *Arcobacter butzleri* strains and comparison with their MLST and PFGE types

Federica Giacometti[a], Silvia Piva[a], Katleen Vranckx[b], Katrien De Bruyne[b], Ilenia Drigo[c], Alex Lucchi[d], Gerardo Manfreda[d], Andrea Serraino[a]

[a]Department of Veterinary Medical Sciences, University of Bologna, Ozzano dell'Emilia, Italy

[b]Applied Maths NV, Sint-Martens-Latem, Belgium

[c]Istituto Zooprofilattico Sperimentale delle Venezie, Villorba, Treviso, Italy

[d]Department of Agricultural and Food Sciences, University of Bologna, Ozzano dell'Emilia, Italy

Corresponding author:

Silvia Piva, Department of Veterinary Medical Sciences, University of Bologna

Via Tolara di Sopra 50, 40064 Ozzano dell'Emilia (BO), Italy

Email: silvia.piva@unibo.it

federica.giacometti3@unibo.it

Katleen_Vranckx@applied-maths.com

katrien_debruyne@applied-maths.com

idrigo@izsvenezie.it

alex.lucchi3@unibo.it

gerardo.manfreda@unibo.it

andrea.serraino@unibo.it

Abstract

For the first time, this study evaluated the use of MALDI-TOF as a typing tool for *Arcobacter butzleri*. A total of 103 *A. butzleri* strains isolated from different sources in an artisanal dairy plant in Italy were identified and typed using MALDI-TOF and compared with their multilocus sequence typing (MLST) and pulsed field gel electrophoresis (PFGE) profiles found in previous studies. MALDI-TOF correctly identified all the isolates to species level. No clearly delineated clusters appeared on dendrograms based on either the complete spectra or the significant peaks, but nine clusters were defined using the cophenetic correlation. Interestingly, MALDI-TOF proved able to discriminate *A. butzleri* strains below species level, confirming its potential use for epidemiological surveys. As expected, the comparative analysis with PFGE and MLST showed that the discriminatory index was lower for MALDI-TOF but roughly comparable to sequence types and pulsotypes. MALDI-TOF appears to be a relatively low cost answer to the urgent need for more rapid, less expensive typing tools suitable for source attribution studies, readily allowing multiple typing methods to be combined. This study provides insights into MALDI-TOF as potential epidemiological tool. Its application in healthcare surveillance systems awaits further exploration to encourage interaction and convergence studies between primary care in humans and animal and food veterinary authorities as part of the One Health concept.

Key words: MALDI-TOF MS, subtyping, *Arcobacter butzleri,* PFGE, MLST.

1. Introduction

The genus *Arcobacter* is an unusual taxon within the epsilon subdivision of Proteobacteria containing both pathogenic and free-living species found in a wide range of environments (Miller *et al.*, 2007). It has long been considered an emerging human enteric pathogen linked to gastrointestinal illnesses (Collado and Figueras, 2011; Hsu and Lee, 2015). Although several aspects of *Arcobacter* epidemiology and virulence are starting to be clarified, key reservoirs and mechanisms of transmission have yet to be fully determined (Collado and Figueras, 2011). *Arcobacter* species are ubiquitous in animals, in a variety of foods of animal and non-animal origin, and in both aquatic and food-processing environments (Collado and Figueras, 2011; Merga *et al.*, 2013), usually showing a high genotype diversity in all these sources. *Arcobacter butzleri* is the best characterized of all *Arcobacter* species. It is probably an environmental organism (Miller *et al.*, 2007) with some level of niche adaptation (Merga *et al.*, 2013) and with the ability to survive in the adverse conditions imposed by food processing and storage (Collado and Figueras, 2011; Ferreira *et al.*, 2015; Giacometti *et al.*, 2013; Giacometti *et al.*, 2015; Hausdorf *et al.*, 2013; Rasmussen *et al.,* 2013; Scarano *et al.*, 2014; Serraino and Giacometti, 2014; Shah *et al.*, 2013) that may cause disease through ingestion of contaminated water or food (Collado and Figueras, 2011; Miller *et al.*, 2007).

Source-attribution studies for the burden of human illness require bacterial typing to identify sources and routes of product contamination. Bacterial typing is also a prerequisite for targeted control measures (Dieckmann *et al.*, 2016) and for source-tracking studies to determine the origin of a specific strain by grouping the sources (Santos *et al.*, 2016). The term subtyping refers to characterization beyond the species or subspecies level, allowing the determination of clonal relationships and the phylogenetic relatedness of bacterial strains (Dieckmann *et al.*, 2016). Nowadays, the genotyping methods most commonly used are based on DNA banding patterns, such as pulsed field gel electrophoresis (PFGE) and amplified fragment length

polymorphism (AFLP), PCR-restriction fragment-length polymorphism (RFLP), random amplification of polymorphic DNA (RAPD), enterobacterial repetitive intergenic consensus (ERIC-PCR), multiple locus variable number of tandem repeats analysis (MLVA), multilocus sequence typing (MLST) and 16S rRNA gene sequencing. All these techniques possess different discriminatory powers, and their use depends on the main objective to be achieved. In spite of their recognized resolution, many of these approaches often lack reproducibility within and among laboratories, whereas others are discriminatory and reproducible but expensive, laborious and time-consuming - all undesirable factors for the identification of contamination sources (Santos *et al.*, 2016).

Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) has wrought the most radical change in the diagnostic microbiology workflow in the last decade (Fournier *et al.*, 2013) and has become a routine tool for microorganism identification in clinical microbiology laboratories worldwide. However, beyond microbe identification, whose importance for human health care is unquestionable, MALDI-TOF MS has proved to have great potential for epidemiological strain typing and antimicrobial susceptibility/resistance detection (Sanguinetti and Posteraro, 2016). This phenotyping technique is based on the detection of a large number of spectral features originating from proteins, namely highly abundant ribosomal and nucleic acid-binding proteins. Though several attempts have been made to apply MALDI-TOF MS to higher resolution microbial discrimination, they have not yielded uniform success, and the limits of the taxonomic resolution of MALDI-TOF MS profiling might be determined in large part by the nature of the particular bacterium profiled (Ghyselinck *et al.*, 2011; Sandrin *et al.*, 2013). Hence, both the taxonomic resolution of MALDI-TOF MS and whether MALDI-TOF MS analysis will overlap other subtyping techniques need to be evaluated individually for a particular genus or species of interest. No such studies have hitherto been performed on *A. butzleri* isolates.

The aim of the present study was to evaluate the ability of MALDI-TOF technology to characterize *A. butzleri* isolates according to their different pattern of TOF peaks, and to perform a comparative analysis of their previously obtained MLST and PFGE profiles (De Cesare *et al.*, 2015; Giacometti *et al.*, 2013).

2. Material and methods

2.1 Strains tested

A set of 103 *A. butzleri* strains, of which 102 were collected from different sources in an artisanal dairy plant in four samplings in the Emilia Romagna Region between October and December 2012, and the references strains *A. butzleri* DSM 8739^T and *A. cryaerophilus* DSM 7289^T previously characterized by pulsed-field gel electrophoresis (PFGE) (Giacometti *et al.*, 2013) and multilocus sequence typing (MLST) (De Cesare *et al.*, 2015) were selected and analysed. Overall, the strains were obtained from food samples (i.e. raw cow and buffalo milk and ricotta cheese) (N = 28); food contact surfaces (i.e. bulk tank valve, cheese vat, drainage table, milk pump and mozzarella molding roller) (N = 45) and non-food contact surfaces (i.e. floors of cooler room and floor drain) (N = 29). The strain details were previously described (Giacometti *et al.*, 2013). Extrapolating the results from the cited studies, the strains collated and used for this study showed a high diversity of 34 PFGE profiles and 21 sequence types (STs) respectively. Table 1 briefly reports the characteristics of the 102 *A. butzleri* strains included in this study and their MLST and PFGE results.

2.2 Sample preparation

Strains were cultured on nutrient agar (Oxoid, Basingstoke, United Kingdom) supplemented with 5% laked horse blood (Oxoid, Basingstoke, United Kingdom), incubated at 30°C for 24 h and subjected to ethanol-formic acid extraction according to the MALDI Biotyper protocol (Bruker Daltonics, Bremen, Germany). Briefly, nuclease-free water (200 µL) was aliquoted

into a plasticizer-free 1.5-ml Eppendorf tube, several loops of organisms from a whole plate were added, and the water and organisms were thoroughly mixed by vortexing. To this suspension of organisms, 900 µL 100% ethanol was added, and again, the organisms and the suspension were thoroughly mixed by vortexing. To deposit the biological material, the tubes were centrifuged at 20,000xg for 2 min and the supernatant was decanted. To remove the residual ethanol, the tubes were centrifuged again and the ethanol was removed by careful pipetting. Twenty microliters of 70% formic acid (Sigma) were added to the pellet, and the formic acid and pellet were well mixed by vortexing, followed by the addition of 20 µL of pure acetonitrile (Sigma-Aldrich), which was carefully mixed with the other components by pipetting up and down. Finally, the tubes were centrifuged at 20,000xg for 2 min, resulting in a supernatant ready for analysis. One microliter of sample was carefully placed on a 96-spot polished steel target (MSP 96 target, ground steel; Bruker Daltonics) with 1 µL of matrix solution, a saturated solution of α-cyano-4-hydroxycinnamic acid (HCAA) matrix (Bruker Daltonics) in 50% acetonitrile, 47.5% water and 2.5% trifluoroacetic acid (Sigma).

MALDI-TOF MS was performed on the MALDI Biotyper Microflex LT controlled by FlexControl software (version 3.3; Bruker Daltonics) at the Istituto Zooprofilattico Sperimentale delle Venezie, Villorba, Treviso, Italy. Spectra were acquired using automatic mode and default settings (2,000 to 20,000 Da; linear positive method); 12 technical replicates for each of 102 strain were generated and we implemented reference *A. butzleri* DSM 8739^T and *A. cryaerophilus* DSM 7289^T for the reference spectra. The raw MALDI spectra database created was then exported and imported in BioNumerics 7.6.1 (Applied Maths NV, Sint-Martens-Latem, Belgium) for the following data analysis.

2.3 Data analysis

Upon import, the background of the spectra was removed using the rolling disk method, the noise was calculated using the continuous wavelet transform (CWT) method and the spectra

were smoothed using a Kaiser window (Monchamp *et al.*, 2007). Peaks were detected in the spectra using the CWT method with a signal-to-noise threshold of 2. The spectra of all technical replicates were summarized to create an average spectrum per isolate, replicates with a correlation to the average of less than 95% were removed and the final average spectrum was calculated with only the remaining replicates. Only peaks present in 75% of the replicates were considered a peak on the average spectrum.

The resulting average spectra were used to calculate a UPGMA dendrogram using a Pearson similarity coefficient. A peak matching was performed with a position tolerance of 500ppm x m/z + 2Da. Peaks from different spectra within this tolerance were considered to belong to the same peak class. As the majority of peaks in this m/z range result from ribosomal proteins that show little variation within a species, a dendrogram was also constructed using only specific subsets of peaks.

The resulting peak matching tables were used to perform Principal Component Analysis (PCA) and Linear Discriminant analysis (LDA) after average intensity-based normalization. Additional statistical analysis was done, using Kruskal-Wallis to test significant differences in peak intensities between different MLST types.

3. Results

MALDI-TOF MS correctly identified all 103 *A. butzleri* strains and 1 *A. cryaerphilus* strain to species level with score values ≥1.9 using the BRUKER BIOTYPER software.

The dendrogram of all the investigated spectra based on the complete spectra (Fig. 1.) revealed no clearly delineated clusters. Most peaks showed no significant differences among the strains tested, with differences caused only by a few peaks. The following peaks were significant at the Kruskal Wallis test (m/z 2940.78; 3753.27; 4215.49; 4349.78; 4363.08; 4413.78; 4420.53; 8698.71; 8724.88; 8827.09; 8840.86; 8866.56; 9419.41; see Fig. 2). A dendrogram based on

these peaks alone (see Fig. 3) showed well-delineated clusters of which nine were defined using the cophenetic correlation to decide whether a cluster was well separated from the other strains. In addition, in Fig. 3, a description of which significant peaks are distributed among the MALDI-groups is reproduced in a heat map revealing peaks intensity and the fact that significant peaks are indeed shared among groups, but each group has a specific combination of peaks .

Interestingly, MALDI TOF was able to recognize and reflect the complex heterogeneity of the food and environmental *A. butzleri* strains considered in this study, and the subtyping results obtained appeared roughly similar to those observed with MLST and PFGE. Simpson's index of diversity (DI) produces a single numerical value (from 0 to 1) to compare the abilities of single or combined typing schemes to discriminate between unrelated isolates and was 0.863, 0.920 and 0.933 for MALDI TOF, MLST and PFGE respectively.

Next, PCA and LDA were applied to distinguish the overall variation among strains and among the MALDI-groups. On the one hand, a PCA based on the complete peak matching table showed no clearly defined groups and no separation of the subtypes obtained with MLST and PFGE (data not shown), similar to the findings reported dendrogran in Fig. 1. On the other, LDA using the ST as groups and only the significantly different peak classes among these STs disclosed several groups (see Fig. 4). Strains from the same ST or pulsotypes (if lower) could be found in the same MALDI cluster, even though some STs and pulsotypes were mixed together.

4. Discussion

Microbiological monitoring of food products and the efficiency of early warning systems and outbreak investigations depend on the rapid identification and strain characterization of pathogens posing risks to the health and safety of consumers (Dieckmann *et al.*, 2016).

227 Although pathogen detection is the first stage of identifying problem areas in a food processing
228 environment or health care system, strain level subtyping is crucial to highlight genotypic
229 differences among strains with particular niches and to trace the sources of contamination in a
230 food processing plant during an outbreak or extended epidemiological investigations, or to track
231 foodborne pathogens throughout production to determine where they enter into the system.
232 Specifically, accurate identification of *A. butzleri* transmission pathways among animals and
233 environmental sources and accurate source tracking studies remain scant. In addition, *A.*
234 *butzleri* biology is not closely associated with any particular host or hosts (Miller *et al.*, 2007).
235 Nor does it possess species-wide pathogenicity, suggesting that *A. butzleri* is an environmental
236 microorganism demanding rapid identification of the contamination source to understand the
237 epidemiology of a disease and the relative contribution of reservoirs, pathways, exposure and
238 risk factors. To obtain these data, a more extensive and rigorous application of high throughput
239 subtyping methods is a priority. Since, MALDI TOF is a rapid low-cost epidemiological
240 method, this study tested its use for the first time as a typing tool for *A. butzleri*.
241 Contrary to species-level identification, strain-level discrimination using MALDI-TOF MS is
242 expected to be influenced by minor changes in the mass spectra (Kern *et al.,* 2014). Therefore,
243 the same culture conditions, namely culture media, culturing time and conditions, were applied
244 to all the investigated samples to avoid the differences observed in the literature (Jadhav *et al.*,
245 2015; Sandrin *et al.*, 2013). Overall, MALDI TOF was able to discriminate the investigated *A.*
246 *butzleri* strains below species level, confirming its potential use for epidemiological surveys.
247 As expected, the comparison analysis with PFGE and MLST showed the DI was lower for
248 MALDI TOF but, at the same time, roughly comparable to ST types and pulsotypes indicating
249 a fair concordance among the typing method results. MALDI TOF was able to trace a possible
250 transmission route from environment to product, or vice versa, within an artisanal dairy plant,
251 reflecting similar results and typing ability previously observed by both MLST and PFGE.

PFGE is generally regarded as "the gold standard" for subspecies classification of microorganisms, and obviously a similar strain-level discrimination and demarcation between food and environmental isolates is difficult to obtain with MALDI TOF. A well-reported but important consideration is that genotypic and phenotyic typing techniques measure completely different cellular properties (Jadhav *et al.*, 2015): MALDI TOF detects mainly conserved ribosomal proteins and, to a lesser extent, nucleic acid binding proteins that exhibit only limited strain specificity (Dieckmann *et al.*, 2016). At the same time, the same set of genetic events relevant for PFGE will give rise to changes in the mobility of proteins during MS and a variable peak number/PFGE restriction fragment number ratio could be tolerated to define related types by MALDI TOF typing (Spinali *et al.*, 2015). MLST is based on allelic nucleic acid variants among housekeeping genes located in the genome of the tested microorganism and has a high discriminatory power with the advantage of providing data readily transportable from one laboratory to another and suitable for global and long-term or evolutionary studies rather than local epidemiology (Dieckmann *et al.*, 2016). On the contrary, an added advantage of MALDI TOF is that the peak intensities which indirectly convey the extent of protein expression are also used to discriminate the isolates, whereas most genotypic methods test only the presence or absence of particular genes (Taneja *et al.*, 2016). Moreover, identification and typing can both be done simultaneously using MALDI TOF with the direct colony plating method (Taneja *et al.*, 2016). Finally, both these conventional typing methods (MLST and PFGE) remain either very labor intensive or very expensive, whereas MALDI TOF provides a low cost per sample analysis and rapid results despite the high initial acquisition costs.

Obviously MALDI TOF has some limitations, namely those linked to different sample preparation and culture conditions which can influence the reproducibility of the method. In addition, MALDI TOF is based on the expression of intrinsic proteins which are conserved and do not show a high degree of evolution, meaning that evolutionary changes may not be

277 reproduced using this technique (Taneja *et al.*, 2016). Further, rigorous statistical analysis is
278 inevitably required to analyse MALDI TOF profiles.
279 Our findings suggest that MALDI TOF may allow multiple typing methods to be combined
280 when certain strains are associated in an outbreak or for epidemiological studies of emergent
281 pathogens like *A. butzleri* that present prohibitive sampling, isolation and genotyping costs if
282 the bacterium is not already integrated in existing surveillance or monitoring programs. For this
283 aspect, typing is likely restricted to relative small cohorts of isolates in either clinical
284 diagnostics or a healthcare setting (Sanguinetti and Posteraro, 2016) and MALDI TOF appears
285 to be a relatively low cost answer to the urgent need for more rapid, less expensive typing tools
286 suitable for source attribution. In addition, faster methods are more realistic than prolonged
287 genotypic methods in an outbreak setting (Taneja *et al.*, 2016), and a standardized MALDI TOF
288 scheme could serve to generate a preliminary cluster of isolates for a result-based isolate
289 selection (Jadhav *et al.*, 2015) to decide if further and more appropriate typing methods are
290 necessary. Similar promising findings were reported for emetic and non-emetic *Bacillus cereus*
291 group members (Sato *et al*., 2017; Fiedoruk *et al.*, 2016) and for *Escherichia coli* (Tagg *et al*.,
292 2015; Christner *et al*., 2014) confirming that the use of MALDI TOF as an automated tool for
293 large-scale populations analyses or for targeted screening seems to be effective.
294 In the case of outbreaks, clinical microbiology laboratories can detect the emergence of
295 unknown species, particular pathotypes or antibiotic resistance patterns, and therefore play an
296 important part in warning the medical authorities, which can corroborate the results across a
297 particular region, country or continent (Fournier *et al.*, 2013). In this context, the potential of
298 MALDI TOF in the field of healthcare surveillance systems merits further exploration and
299 validation.
300 The acquisition of MALDI TOF MS is recommended for leading microbiological laboratories
301 where its daily use for bacterial identification and a systematic elaboration of MS peaks for

epidemiological strain typing and antimicrobial susceptibility/resistance detection would be useful to create a shared database for strains of interest. The exchange of data between human and animal/food care surveillance systems could serve to organize a standardized central warning system and reflects the need for convergence of these domains under the One Health concept.

Conflict of Interest

Katleen Vranckx and Katrien De Bruyne are employees of Applied Maths NV.

Acknowledgments. This research did not receive specific grant from any funding agencies in the public, commercial, or not-for-profit sectors. Anne Collins edited the English manuscript.

References

Collado, L., Figueras, M.J., 2011. Taxonomy, epidemiology, and clinical relevance of the genus Arcobacter. Clin. Microbiol. Rev. 24, 174-192.

Christner, M., Trusch, M., Rohde, H., Kwiatkowski, M., Schlüter, H., Wolters, M., Aepfelbacher, M., Hentschke, M., 2014. Rapid MALDI-TOF mass spectrometry strain typing during a large outbreak of Shiga-Toxigenic *Escherichia coli*. PLoS One. 8;9(7):e101924.

De Cesare, A., Parisi, A., Giacometti, F., Serraino, A., Piva, S., Caruso, M., De Santis, E.P.L., Manfreda, G., 2015. Multilocus sequence typing of Arcobacter butzleri isolates collected from dairy plants and their products, and comparison with their PFGE types. J. Appl. Microbiol. 120(1), 165-174.

Dieckmann, R., Hammerl, J.A., Hahmann, H., Wicke, A., Kleta, S., Dabrowski, P.W., Nitsche, A., Stämmler, M., Al Dahouk, S., Lasch, P., 2016. Rapid characterisation of Klebsiella oxytoca isolates from contaminated liquid hand soap using mass spectrometry, FTIR and Raman

spectroscopy. Faraday Discuss. 187, 353–375.

Ferreira, S., Queiroz, J.A., Oleastro, M., Domingues, F.C., 2015. Insights in the pathogenesis and resistance of Arcobacter: A review. Crit. Rev. Microbiol. 42(3), 364-383.

Fiedoruk, K., Daniluk, T., Fiodor, A., Drewicka, E., Buczynska, K., Leszczynska, K., Bideshi, D.K., Swiecicka, I., 2016. MALDI-TOF MS portrait of emetic and non-emetic *Bacillus cereus* group members. Electrophoresis. 37(15-16):2235-47.

Fournier, P.E., Drancourt, M., Colson, P., Rolain, J.M., La Scola, B., Raoult, D., 2013. Modern clinical microbiology: new challenges and solutions. Nat. Rev. Microbiol. 11(8), 574-585.

Ghyselinck, J., Van Hoorde, K., Hoste, B., Heylen, K., De Vos, P., 2011. Evalua- tion of MALDI-TOF MS as a tool for high-throughput dereplication. J. Microbiol. Methods. 86, 327–336.

Giacometti, F., Lucchi, A., Di Francesco, A., Delogu, M., Grilli, E., Guarniero, I., Stancampiano, L., Manfreda, G., Merialdi, G., Serraino, A., 2015. Arcobacter butzleri, Arcobacter cryaerophilus, and Arcobacter skirrowii Circulation in a Dairy Farm and Sources of Milk Contamination. Appl. Environ. Microbiol. 81(15), 5055-5063.

Giacometti, F., Lucchi, A., Manfreda, G., Florio, D., Zanoni, R.G., Serraino, A., 2013. Occurrence and genetic diversity of Arcobacter butzleri in an artisanal dairy plant in Italy. Appl. Environ. Microbiol. 79(21), 6665-6669.

Hausdorf, L., Neumann, M., Bergmann, I., Sobiella, K., Mundt, K., Fröhling, A., Schlüter, O., Klocke, M., 2013. Occurrence and genetic diversity of Arcobacter spp. in a spinach-processing plant and evaluation of two Arcobacter-specific quantitative PCR assays. Syst. Appl. Microbiol. 36(4), 235-243.

Hsu, T.T., Lee, J., 2015. Global distribution and prevalence of Arcobacter in food and water. Zoonoses Public Health. 62(8), 579-589.

Jadhav, S., Gulati, V., Fox, E.M., Karpe, A., Beale, D.J., Sevior, D., Bhave, M., Palombo, E.A.,

352 2015. Rapid identification and source-tracking of Listeria monocytogenes using MALDI-TOF
353 mass spectrometry. Int. J. Food Microbiol. 202, 1-9.
354 Kern, C.C., Vogel, R.F., Behr, J., 2014. Differentiation of Lactobacillus brevis strains using
355 matrix-assisted-laser-desorption-ionization-time-of-flight mass spectrometry with respect to
356 their beer spoilage potential. Food Microbiol. 40, 18-24.
357 Merga, J.Y., Winstanley, C., Williams, N.J., Yee, E., Miller, W.G., 2013. Complete genome
358 sequence of the Arcobacter butzleri cattle isolate 7h1h. Genome Announc. 1(4), e00655-13.
359 Miller, W.G., Parker, C.T., Rubenfield, M., Mendz, G.L., Wösten, M.M., Ussery, D.W., Stolz,
360 J.K., Binnewies, T.T., Hallin, P.F., Wang, G., Malek, J.A., Rogosin, A., Stanker, L.H., Malek,
361 J.A., 2007. The complete genome sequence and analysis of the Epsilonproteobacterium
362 Arcobacter butzleri. PLoS One. 2(12), e1358.
363 Monchamp, P., Andrade-Cetto, L., Zhang, J., Henson, R., 2007. Signal processing methods for
364 mass spectrometry. In: Alterovitz, G., Ramoni, M. (eds), Systems Bioinformatics: An
365 Engineering Case-Based Approach. Artech House inc., Norwood, MA, pp. 101-124.
366 Rasmussen, L.H., Kjeldgaard, J., Christensen, J.P., Ingmer, H., 2013. Multilocus sequence
367 typing and biocide tolerance of Arcobacter butzleri from Danish broiler carcasses. BMC Res.
368 Notes. 6(1), 322.
369 Sandrin, T.R., Goldstein, J.E., Schumaker, S., 2013. MALDI TOF MS profiling of bacteria at
370 the strain level: a review. Mass. Spectrom. Rev. 32(3), 188-217.
371 Sanguinetti, M., Posteraro, B., 2016. Mass spectrometry applications in microbiology beyond
372 microbe identification: progress and potential. Expert. Rev. Proteomics. 13(10), 965-977.
373 Santos, I.C., Hildenbrand, Z.L., Schug, K.A., 2016. Applications of MALDI-TOF MS in
374 environmental microbiology. Analyst. 141(10), 2827-2837.
375 Sato, J., Nakayama, M., Tomita, A., Sonoda, T., Hasumi, M., Miyamoto, T., 2017. Evaluation
376 of repetitive-PCR and matrix-assisted laser desorption ionization-time of flight mass

spectrometry (MALDI-TOF MS) for rapid strain typing of *Bacillus coagulans*. PLoS One. 12(10):e0186327.

Scarano, C., Giacometti, F., Manfreda, G., Lucchi, A., Pes, E., Spanu, C., De Santis, E.P.L., Serraino, A., 2014. Arcobacter butzleri in sheep ricotta cheese at retail and related sources of contamination in an industrial dairy plant. Appl. Environ. Microbiol. 80(22), 7036-7041.

Serraino, A., Giacometti, F., 2014. Short communication: Occurrence of Arcobacter species in industrial dairy plants. J. Dairy Sci. 97(4), 2061-2065.

Shah, A.H., Saleha, A.A., Zunitam Z., Murugaiyah, M., Aliyu, A.B., Jafri, N., 2013. Prevalence, distribution and antibiotic resistance of emergent Arcobacter spp. from clinically healthy cattle and goats. Transbound. Emerg. Dis. 60(1), 9-16.

Spinali, S., van Belkum, A., Goering, R.V., Girard, V., Welker, M., Van Nuenen, M., Pincus, D.H., Arsac, M., Durand, G., 2015. Microbial Typing by Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry: Do We Need Guidance for Data Interpretation? J. Clin. Microbiol. 53(3), 760-765.

Tagg, K.A., Ginn, A.N., Partridge, S.R., Iredell, J.R., 2015. MALDI-TOF Mass Spectrometry for Multilocus Sequence Typing of *Escherichia coli* Reveals Diversity among Isolates Carrying *bla*CMY-2-Like Genes. PLoS One. 10(11): e0143446.

Taneja, N., Sethuraman, N., Mishra, A., Mohan, B., 2016. The 2002 Chandigarh cholera outbreak revisited: utility of MALDI-TOF as a molecular epidemiology tool. Lett. Appl. Microbiol. 62(6), 452-458.

Table 1 Typology, source, sampling time, multilocus sequence typing and pulsed-field gel

402 electrophoresis results of the 102 *A. butzleri* strains included in this study

Strain	Source	Sampling	ST	Pulsotype
	Food isolates			
56	Raw cow milk	II	437	1
45	Raw cow milk	III	423	2
162	Raw cow milk	III	423	1
210	Raw cow milk	IV	437	7
224	Raw cow milk	IV	438	6
127	Raw WB milk	I	420	33
132	Raw WB milk	I	420	32
149	Raw WB milk	I	432	12
68	Raw WB milk	II	425	26
77	Raw WB milk	II	425	26
91	Raw WB milk	II	425	26
95	Raw WB milk	II	428	26
35	Raw WB milk	III	435	25
37	Raw WB milk	III	422	5
40	Raw WB milk	III	423	2
42	Raw WB milk	III	424	4
108	Raw WB milk	III	429	28
39	Raw WB milk	IV	422	5
185	Raw WB milk	IV	427	10
207	Raw WB milk	IV	436	34
219	Raw WB milk	IV	436	14
220	Raw WB milk	IV	436	14
71	Ricotta cheese	II	66	11
101	Ricotta cheese	II	66	11
117	Ricotta cheese	II	66	11
163	Ricotta cheese	II	66	9
198	Ricotta cheese	II	66	11
205	Ricotta cheese	II	66	29
	Isolates from food contact surfaces			
93	Bulk tank valve	I	420	33
124	Bulk tank valve	II	430	16
125	Bulk tank valve	II	429	28
109	Bulk tank valve	III	430	16
172	Bulk tank valve	III	434	17
183	Bulk tank valve	IV	430	16
190	Bulk tank valve	IV	430	20
92	Cheese vat	I	427	9
114	Cheese vat	I	427	8
123	Cheese vat	I	419	4
99	Cheese vat	II	421	22
120	Cheese vat	II	431	22
121	Cheese vat	II	435	25

20	Cheese vat	III	434	19
21	Cheese vat	III	434	19
197	Cheese vat	IV	430	16
199	Cheese vat	IV	430	16
239	Cheese vat	IV	430	18
7	Drainage table	III	435	25
10	Drainage table	III	434	15
11	Drainage table	III	435	22
201	Drainage table	IV	435	25
216	Drainage table	IV	435	25
217	Drainage table	IV	435	25
113	Milk pump	I	421	22
159	Milk pump	I	433	20
3	Milk pump	II	419	25
58	Milk pump	II	419	22
4	Milk pump	III	419	21
49	Milk pump	III	435	25
60	Milk pump	III	419	23
64	Milk pump	III	420	32
86	Milk pump	III	438	24
158	Milk pump	IV	427	8
187	Milk pump	IV	427	9
195	Milk pump	IV	427	8
196	Milk pump	IV	427	9
106	Mozzarella cheese molding roller	I	421	24
46	Mozzarella cheese molding roller	II	435	27
47	Mozzarella cheese molding roller	II	435	25
62	Mozzarella cheese molding roller	II	420	32
12	Mozzarella Cheese molding roller	III	420	30
48	Mozzarella cheese molding roller	III	435	25
152	Mozzarella cheese molding roller	IV	432	13
240	Mozzarella cheese molding roller	IV	435	25
	Isolates from nonfood contact surfaces			
57	Cooler room floor	II	424	4
67	Cooler room floor	II	435	25
74	Cooler room floor	II	424	4
78	Cooler room floor	II	424	4
103	Cooler room floor	II	424	4
105	Cooler room floor	II	420	33
112	Cooler room floor	II	429	25
115	Cooler room floor	II	435	25
116	Cooler room floor	II	424	4
25	Cooler room floor	III	421	24
27	Cooler room floor	III	420	33
30	Cooler room floor	III	435	25
31	Cooler room floor	III	420	33

32	Cooler room floor	III	420	33
111	Floor drain	I	420	33
135	Floor drain	I	438	3
80	Floor drain	II	426	3
81	Floor drain	II	435	25
107	Floor drain	II	421	22
118	Floor drain	II	435	25
15	Floor drain	III	435	25
16	Floor drain	III	422	25
17	Floor drain	III	435	25
18	Floor drain	III	435	25
19	Floor drain	III	435	25
43	Floor drain	III	423	1
177	Floor drain	IV	427	31
178	Floor drain	IV	420	33
179	Floor drain	IV	427	9

Fig. 1. MSP dendrogram of all the 102 *A. butzleri* strains based on the complete spectra compared with their MLST and PFGE types. Different colored squares represent MLST and PFGE types.

Fig. 2. Two representative *A. butzleri* mass-spectral profiles showing MALDI TOF peaks differing significantly between different MLST profiles, at *m/z* 4413.8, *m/z* 4420.5 and at *m/z* 8827.1, *m/z* 8840.9 and *m/z* 8866.5; x-axis contains *m(Da)/z*, y-axis intensity with left the absolute intensity in arbitrary units and right the relative intensity compared to the most intense peak in the profile.

Fig. 3. MSP dendrogram of the *A. butzleri* strains based on the peaks significantly different STs compared with their MLST and PFGE types with the MALDI-group marked on the right. Peak intensity is represented using different colors ranges from blue (low intensity) to green and yellow (medium intensity) and red (high intensity).

Fig. 4. Linear discriminant analysis with isolates coloured according to their sequence type shows the majority of isolates fall into the same group with isolates of the same ST, though several STs cannot be separated from each other. The data is visualized with the three components containing the highest discrimination between the STs, the x-axis contributes to 43.9 % of discrimination, the y-axis for 22.3% and the z-axis for 18.2%.