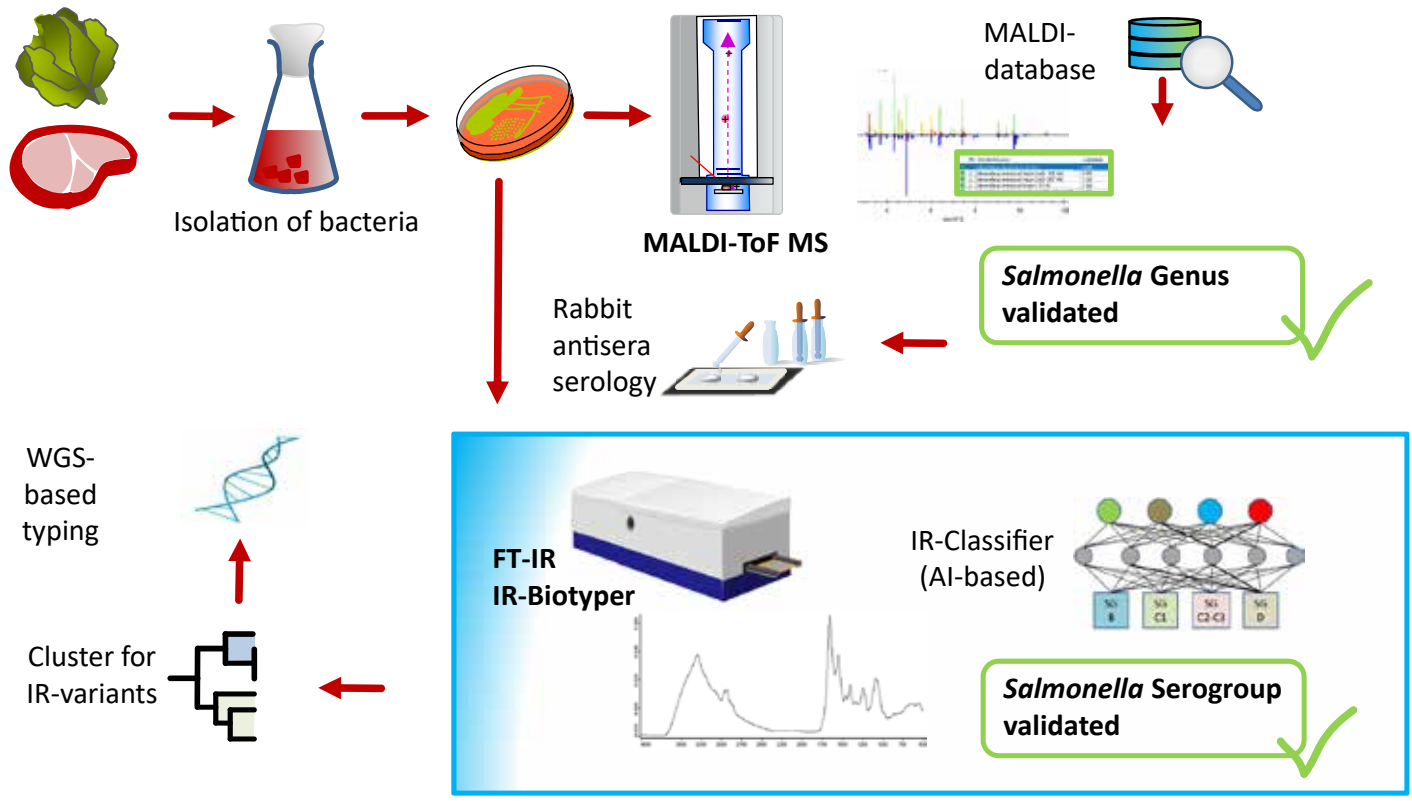
cvua STUTTGART
Chemisches und Veterinäruntersuchungsamt Stuttgart

Aspects of Food Control and Animal Health

eJournal 2025, Volume 16



Establishment and Thorough External Validation of an FTIR Spectroscopy Classifier for *Salmonella* Serogroup Differentiation

Helene Oberreuter, Miriam Cordovana, Martin Dyk, Jörg Rau

Impressum

Aspects of Food Control and Animal Health
August 2025, Volume 16

Publischer
Baden-Württemberg
Chemisches und Veterinäruntersuchungsamt
Stuttgart
70702 Fellbach, P.O. Box: 1206, Germany
Phone: +49 711 3426 1234
Email: postfach@cvuas.bwl.de
Internet: www.cvuas.de

Responsible according to the German Press Law:
Nadja Bauer (Director of CVUA Stuttgart)

Edited by
CVUA Stuttgart Public Relation Unit

Cover picture
»Workflow of *Salmonella* identification« by Jörg Rau, CVUA Stuttgart

Layout, graphics and image editing:
Pat Schreiter, CVUA Stuttgart

ISSN: 2196-3460
https://doi.org/10.48414/aspects2025/16

Aspects of Food Control and Animal Health is a free e-journal of CVUA Stuttgart.
It is dedicated to the publication of analytical methods, results, and case reports from CVUA Stuttgart and, per invitation, from other institutions that are of broad interest to scientists and users engaged in food control and animal health.

Establishment and Thorough External Validation of an FTIR Spectroscopy Classifier for *Salmonella* Serogroup Differentiation

Helene Oberreuter[1]*, Miriam Cordovana[2], Martin Dyk[1] & Jörg Rau[1]

[1] Chemisches und Veterinäruntersuchungsamt Stuttgart (CVUAS), 70736 Fellbach, Germany
[2] Bruker Daltonics GmbH & Co. KG, Bremen, Germany

Article Info

Keywords:
Salmonella serogroup differentiation, FTIR spectroscopy, classifier validation

Supplementary material:
Suppl. 1, Suppl. 2, Suppl. 3 and Suppl. 4 available at https://doi.org/10.48414/aspects2025/16

Article history:
Preprint: 18 March 2025 in BioRxiv™
Available online: https://doi.org/10.1101/2025.03.18.643663

Abstract

As one of the most relevant food-borne pathogens, the reliable detection, confirmation and fine-typing of *Salmonella* strains is very important. *Salmonella* serotype determination using rabbit antisera is the worldwide-accepted standard, but it is labor intensive, costly, and requires extensive experience. As an alternative, successful discrimination between strains of different serogroups by FTIR spectroscopy has been previously developed for various bacterial groups. For the current study, firstly an FTIR Classifier operating on an IR Biotyper® spectrometer (Bruker, Germany) was designed to distinguish between *n*=36 different *Salmonella* serogroups. An FTIR classifier is an AI-based tool used in FTIR spectroscopy for the analysis and classification of different materials based on their infrared spectra.

Secondly, the differentiation performance of this classifier was determined by a thorough external single-lab validation carried out inspired by the "Guidelines for Validating Species Identifications Using MALDI-ToF-MS" issued by the German Federal Office of Consumer Protection for a targeted identification. The most common *Salmonella* serogroups in Europe, serogroups O:4 (B), O:6,7 (C1), O:8 (C2-C3) and O:9 (D1) were chosen as target parameters and validated using a total of *n*=1,039 infrared absorbance spectra from a total of *n*=167 strains pertaining to *n*=39 serogroups. In sum, serogroups O:4, O:6,7 and O:9 perfectly met the adapted guideline requirements and resulted in a > 99% inclusivity each. Serogroup O:8 arrived at a 96.1% true-positive rate, due to one deviating strain. This validated classification method can thus be used in routine analysis for quick and easy differentiation of the most common *Salmonella* serogroups in food surveillance. In addition, when using the cluster analysis tools of the IR BT®, preselecting isolates before subjecting them to thorough serotyping decreases the workload in current routine analyses.

1. Introduction

Salmonella is an important food pathogen, with salmonellosis being the second most commonly reported gastrointestinal infection in Europe [1]. Nontyphoidal salmonellosis usually manifests itself by a self-limiting diarrhea possibly accompanied by fever that occurs within 6 to 72 hours after ingestion of the living cells [2]. While salmonellosis outbreaks occurred much more frequently in the past, the practice of vac-

***CORRESPONDING AUTHOR:** Dr. Helene Oberreuter, email: helene.oberreuter@cvuas.bwl.de, tel.: +4971134261249
CITATION: Oberreuter H, Cordovana M, Dyk M, Rau J (2025): Establishment and Thorough External Validation of an FTIR Spectroscopy Classifier for *Salmonella* Serogroup Differentiation, 16, 1–18. https://doi.org/10.48414/aspects2025/16

cinating breeding and laying hens with live strains of *Salmonella* (S.) Enteritidis and *S.* Typhimurium has helped to greatly decrease the frequency of cases in humans [3]. Nevertheless, there are still about 91,000 cases occurring annually in Europe [4] and about 1.3 million in North America [4]. A frequent carrier of *Salmonella* continues to be reptiles, whose colonization with *Salmonella* isn't, however, usually revealed by the presence of any symptoms of illness [5]. Moreover, *Salmonella* bacteria occur widely throughout the environment and are frequently isolated from untreated animal food sources such as meat, eggs and milk in their raw condition. Contaminated fresh produce or spices have also been the sources of outbreaks [2]. Several outbreaks implying contaminated ready-to-eat sprouts have been reported, for instance, where such bacteria germinate just as effectively in a warm and moist growth environment as the plant seedlings themselves [6]. Therefore, a quick and simple identification of presumptive colonies is still of great importance.

At the species level, *Salmonella* strains are taxonomically organized into two species, namely *S. enterica* and *S. bongori*. The former of these has been subdivided into six subspecies: *S. enterica* ssp. *enterica*, ssp. *salamae*, ssp. *arizonae*, ssp. *diarizonae*, ssp. *houtenae* and ssp. *indica* [7]. While this number is still manageable, at the serotype level, more than 2,600 *Salmonella* serovars are recognized currently [8]. Of these, more than 1,500 serovars belong to the subspecies enterica, 99% of which may cause infections in animals and humans [9]. The gold standard for *Salmonella* typing is the serotype determination using rabbit antisera. For any given strain, this fine-typing method results in a 3-unit combination code, starting with the serogroup indicated by the letter O, followed by two flagellar antigen units H1 and H2. To date, the 2,600 currently recognized serovars have been grouped into 46 O-groups [7].

At the CVUAS strain collection, there are more than *n*=2,600 deep frozen *Salmonella* field strains comprising *n*=36 O-groups. They have been isolated from both food products and veterinary samples of domestic, farm and zoo animals. Selected representatives from both the isolate collections at CVUAS and the infrared instrument manufacturer Bruker were first used to establish a Fourier Transform infrared spectroscopy (FTIR) classifier that would enable the allocation of unknown salmonella sample isolates to the respective serogroups: the so-called *Salmonella serogroup classifier v3*. Secondly, a different sample set was used to perform an external validation of the FTIR classifier with respect to its capacity for serogroup discrimination.

FTIR spectroscopy is a physicochemical tool that enables discrimination between different microbial species, subspecies or even serogroups. In principle, dried cells are subjected to infrared light in the mid-infrared region. The light absorption is recorded and the resulting spectra are further treated, e.g. normalized and differentiated to the first or second derivative to enable their further usage. Based on databases consisting of these pre-treated spectra, methods using artificial neural networks can be developed that enable discrimination between the microbial target units. FTIR is a comparatively simple and affordable method, whose suitability for bacterial typing has been demonstrated before (for reviews, see [10 – 13]). Specifically, the distinctions between the highly similar *Salmonella* serogroups and serovars has been investigated previously [14 – 26].

The current study outlines the construction setup of a commercial salmonella IR-Biotyper© (IR BT©) classifier and its thorough and formalized external validation carried out in a single-lab. Firstly, the commercially available classifier *Salmonella serogroup classifier v3* was developed using a support vector machine algorithm included in the IR Biotyper® software which aims to distinguish between *n*=36 salmonella serogroups, including the four serogroups containing the serovars *S.* Typhimurium (O:4), *S.* Infantis (O:6,7), *S.* Newport (O:8) and *S.* Enteritidis (O:9). These are the most commonly occurring food and veterinary *Salmonella* isolates both in Europe [4]and at the CVUAS lab. A further class, named "Others", was also included, comprising strains other than the above-mentioned 36 serogroups as well as strains exhibiting a rough phenotype without any somatic antigens.

Secondly, a structured external single-lab validation of this classifier was carried out inspired by the adapted national Guidelines for Validating Species Identifications using MALDI-ToF MS [27, 28] as previously demonstrated for the differentiation of *Listeria monocytogenes* serogroups [29]. The *Salmonella* classifier was challenged with a set of *n*=167 different strains belonging to *n*=39 different serogroups. In addition, the *Salmonella serogroup classifier v3* was tested with a group of *n*=16 salmonella strains from lab proficiency tests performed in our lab from 2006 until 2024.

2. Materials and Methods

2.1. Strains under investigation including confirmation of species by MALDI-ToF MS

All isolates used for the validation had been originally isolated at CVUA Stuttgart from food and animal samples according to ISO 6579 [30], followed by subsequent MALDI-ToF MS analysis for formal confirmation at the species level. The MALDI-ToF MS analysis was done on the MALDI Biotyper system, consisting of an LT microflex mass spectrometer and Compass software, combined with database version K (all Bruker Daltonics, Bremen, Germany) in accordance with the certified workflow for *Salmonella* species confirmation [31]. Prior to their utilization in the current study, the identification of each of these isolates had been confirmed or determined, respectively, by the National Reference Laboratory (NRL *Salmonella*) at the German Federal Institute for Risk Assessment (BfR), together with their respective serotype information.

The list of isolates used for setting up the commercially available *Salmonella serogroup classifier v3* consisted of *n*=158 isolates from *n*=36 different serogroups. Emphasis was placed on serogroups O:4, O:6,7, O:8 and O:9, from which 75 different strains, i.e. 10-30 each, were included, respectively. The remaining 83 strains were arbitrarily chosen, pertaining to serogroups O:3 and O:11 through O:65. These included rough forms without any positive antisera reaction, in order to comprise a challenging outgroup. The intention was to set up a comparatively large and diverse set of strains as a sort of background noise to challenge the classifier with real-(lab)life isolates. As strains from two different serotypes were arbitrarily selected per serogroup, this outgroup was referred to as the "Noah's Ark" strain set.

Likewise, the list of isolates which were used for the external validation of the *Salmonella serogroup classifier v3* comprised *n*=167 isolates from the same *n*=39 O-groups (Table 1, p. 4). None of these had been used by the manufacturer to establish the classifier. With respect to the most common serogroups O:4, O:6,7, O:8 and O:9, *n*=25 different strains from each group were included in the validation set (see Table 1).

In addition, the *Salmonella serogroup classifier v3* was challenged with a set of *n*=16 salmonella strains from lab proficiency tests performed in our lab from 2006 until 2024 (Table 2, p. 8). This set included *n*=6 strains from serogroup O:4 (B), and *n*=5 strains from each of serogroups O:6,7 (C1) and O:8 (C2-C3).

2.2. Incubation of isolates and IR BT® sample preparation

For the setup of the *Salmonella serogroup classifier v3*, the strains were cultivated on seven different solid media (Columbia agar with 5 % sheep blood, Chocolate agar, Tryptose soy agar, Mueller-Hinton agar, XLD agar, *Salmonella-Shigella* agar and MacConkey agar – Becton Dickinson, Heidelberg, Germany) for 24 ± 2 h at 35 ± 2 °C.

For the validation of the *Salmonella serogroup classifier v3*, the strains (Table 1) were incubated in a confluent lawn on Columbia Sheep-Blood (COL-SB) agar (BD Columbia-Agar with 5 % sheep blood by Becton Dickinson, Heidelberg, Germany) for 24 ± 0.5 h at 37 °C. All preparations were performed as previously outlined [29]. Each isolate was cultivated on at least two different COL-SB agar plates and prepared by two different operators (two biological replicates) to simulate preparatory lab variance. Each biological replicate was measured in technical triplicate, yielding at least six spectra per isolate.

For the proficiency test strain group challenge, the strains (Table 2, p. 8) were incubated as described above by preparing two technical replicates from two biological replicates each, amounting to four spectra per strain.

IR BT® sample preparation was performed according to the manufacturer's instructions using the IR Biotyper kit (Bruker Daltonics) [32]. From each agar plate, using a 1 µl loop, cell material was transferred into a 1.5 ml vial containing 50 µl of a 70 % (v/v) ethanol solution, and subsequently vortexed to yield a suspension (one suspension replicate per biological replicate). Afterwards, 50 µl deionized water was added to each vial and the suspension was mixed by pipetting. Each suspension was transferred onto three spots of a silicon sample plate at 15 µl per spot (three measurement replicates per suspension). After all spots were filled, the sample plate was left to dry within a 37 °C-incubator for approximately 15 min until none of the spots showed any visible humidity.

2.3. Recording and processing of spectra

Spectra absorption was recorded in transmission mode on an IR BT® spectrometer (Bruker Daltonics) between wave numbers 4,000 and 500 cm^{-1}. The sample plate containing the dried suspension spots was inserted into the instrument's measurement chamber that is continuously being purged with dry air. The manufacturer's method of acquisition used a resolu-

Table 1. *Salmonella* (*S.*) strains used for the external validation of the Bruker IR-BT® *Salmonella serogroup classifier v3*

Consec. No.	Strain No.	Serogroup Letter/Name	Serotype	Source
1	35493	*S.* B Abony	(1),4,12,27:b:e,n,x	Wild boar
2	7666.2	*S.* B Abortusovis	4,12:c:1,6	Sheep
3	1075.4	*S.* B Agona	(1),4,12:f,g,s:-	Dog, fecal sample
4	5870	*S.* B Brandenburg	4,12:l,v:e,n,z15	PigletFerkel
5	31295.2	*S.* B Bredeney	(1),4,12:l,v:1,7	Turkey meat
6	32369	*S.* B Chester	(1),4,5,12:e,h:e,n,x	Turkey meat, seasoned
7	30692	*S.* B Coeln	(1),4,5,12:y:1,2	Turkey breast steak
8	34975	*S.* B Derby	(1),4,12:f,g:	Pig, sock swab
9	34997	*S.* B Derby	(1),4,12:f,g:	Pig, sock swab
10	11719	*S.* B Hessarek	4,12:a:1,5	Fox
11	33884	*S.* B Indiana	(1),4,12:z:1,7	Dog, fecal sample
12	33881	*S.* B Paratyphi B, d-tartrate positive	(1),4,5,12:b:1,2	Rock python
13	32051	*S.* B Saintpaul	(1),4,12:e,h:1,2	Turkey, sock swab
14	32797	*S.* B Schleissheim	4,12,27:b:-	Laying hen, sock swab
15	9912.2	*S.* B Schwarzengrund	(1),4,12:d:1,7	Turkey leg
16	34886	*S.* B Typhimurium	(1),4,12:i:1,2	Pigeon, liver
17	34934	*S.* B Typhimurium	(1),4,12:i:1,2	Pig, organ
18	35013	*S.* B Typhimurium	(1),4,5,12:i:1,2	Goose leg
19	9929	*S.* B Typhimurium	1,4,12:i:1,2	Liquid egg
20	34398	*S.* B Typhimurium	(1),4,5,12:i:1,2	Racoon
21	33381	*S.* B Typhimurium	(1),4,5,12:i:1,2	Alpaca, liver
22	33385	*S.* B Typhimurium	(1),4,5,12:i:1,2	Hedgehog
23	35875	*S.* B Typhimurium, monophasic	(1),4,12:i:-	Piglet, fecal sample
24	5601.3	*S.* B Typhimurium, monophasic	(1),4,5,12:i:-	Sheep, lung
25	11067.3	*S.* B Typhimurium, monophasic	(1),4,5,12:i:-	Hedgehog, organ
26	2323.2	*S.* C1 Branderup	6,7:e,h:e,n,z15	Iguana
27	7091.2	*S.* C1 Isangi	6,7:d;1,5	Chicken, raw
28	326.4	*S.* C1 Livingstone	6,7:d:l,w	Radish sprouts
29	35069	*S.* C1 Livingstone	6,7:d:l,w	Pig, sock swab
30	6528.3	*S.* C1 Mbandaka	61:k:1,5,7	Chicken
31	796.4	*S.* C1 Mbandaka	6,7:z10:e,n,z15	Turkey, livestock farming dust
32	10598	*S.* C1 Montevideo	6,7:g,m,s:-	Animal feed
33	10593	*S.* C1 Oakland	6,7:z:1,6	Dry, bitter leaves
34	34935	*S.* C1 Ohio	6,7:b:l,w	Pig, sock swab
35	35075	*S.* C1 Oranienburg	6,7:m,t:-	Snake, fecal sample
36	7232	*S.* C1 Oslo	6,7:a:e,n,x	Rat snake
37	32033	*S.* C1 *ssp. enterica, monophasic*	6,7:-:1,5	Chicken breast
38	35033	*S.* C1 *ssp. houtenae*	6,7:z4,z24:-	Rock python, organ
39	33538	*S.* C1 Tennessee	6,7:z29:-	Rock lizard, fecal sample
40	35226	*S.* C1 Thompson	6,7:k:1,5	Rock python, fecal sample
41	4808.3	*S.* C1 Infantis	6,7:r:1,5	Broiler
42	5138.3	*S.* C1 Infantis	6,7:r:1,5	Pig

Table 1 (cont.). *Salmonella* (*S.*) strains used for the external validation of the Bruker IR-BT® *Salmonella serogroup classifier v3*

Consec. No.	Strain No.	Serogroup Letter/Name	Serotype	Source
43	31253	S. C1 Infantis	6,7:r:1,5	Chicken meat
44	31297	S. C1 Infantis	6,7:r:1,5	Minced meat, seasoned
45	35791	S. C1 Infantis	6,7:r:1,5	Chicken meat
46	11506.3	S. C1 Infantis	6,7:r:1,5	Turkey meat
47	85.2	S. C1 Infantis	6,7:r:1,5	Turkey breast
48	32492	S. C1 Infantis	6,7:r:1,5	Chicken breast
49	32981	S. C1 Infantis	6,7:r:1,5	Chicken breast
50	33075	S. C1 Infantis	6,7:r:1,5	Chicken meat
51	4610	S. C2-C3 Bardo	8:e,h:1,2	Garter snake
52	5004.2	S. C2-C3 Blockley	6,8:k:1,5	Turkey meat
53	9383	S. C2-C3 Bovismorbificans	6,8:r:1,5	Turkey liver
54	7940.2	S. C2-C3 Bovismorbificans	6,8:r:1,5	Sprouts blend
55	4586.2	S. C2-C3 Eboko	6,8:b:1,7	Badger
56	34996	S. C2-C3 Eboko	6,8:b:1,7	Badger, organ
57	86.2	S. C2-C3 Glostrup	6,8:z10:e,n,z15	Pepper
58	6844.2	S. C2-C3 Hadar	6,8:z10:e,n,x	Laying hen, sock swab
59	5453.2	S. C2-C3 Hadar	6,8:z10:e,n,x	poultry
60	4446.2	S. C2-C3 Herston	6,8:d:e,n,z15	Reptile
61	1148.2	S. C2-C3 Kottbus	6,8:e,h:1,5	Turkey meat
62	33265	S. C2-C3 Kottbus	6,8:e,h:1,5	Marten
63	31154	S. C2-C3 Muenchen	6,8:d:1,2	Turkey meat
64	32815	S. C2-C3 Newport	6,8:e,h:1,2	Turkey breast
65	32816	S. C2-C3 Newport	6,8:e,h:1,2	Asthma weed
66	34932	S. C2-C3 Newport	6,8:e,h:1,2	Moringa, rubbed
67	34971	S. C2-C3 *ssp. enterica, monophasic*	6,8:-:1,2	Green anaconda, lung
68	33992	S. C2-C3 Stourbridge	6,8:b:1,6	Laying hen, sock swab
69	31528	S. C2-C3 Virginia	8:d:1,2	Horse
70	9531	S. C2-C3 Kentucky	8,20:i:z6	Turkey wings
71	1070	S. C2-C3 Kentucky	8,20:i:z6	Turkey cutlet
72	11190	S. C2-C3 Kentucky	8,20:i:z6	Turkey meat
73	11201	S. C2-C3 Kentucky	8,20:i:z6	Turkey steak
74	31155	S. C2-C3 Kentucky	8,20:i:z6	Turkey meat
75	31255	S. C2-C3 Kentucky	8,20:i:z6	Turkey
76	11556.2	S. D1 Dublin	(1),9,12:g,p:-	Cow
77	7670.2	S. D1 Dublin	(1),9,12:g,p:-	Calf, fecal sample
78	9544.2	S. D1 Dublin	(1),9,12:g,p:-	Ground beef
79	11434	S. D1 Eastbourne	(1),9,12:e,h:1,5	Reptile, fecal sample
80	30458	S. D1 Eastbourne	(1),9,12:e,h:1,5	Iguana, fecal sample
81	11298.3	S. D1 Gallinarum	(1),9,12:-:-	Chicken liver
82	32489	S. D1 Gallinarum	(1),9,12:-:-	Chicken
83	11284	S. D1 Gallinarum	(1),9,12:-:-	Chicken
84	7511	S. D1 Javiana	9,12:l,z28:1,5	Reptile

Table 1 (cont.). *Salmonella* (*S.*) strains used for the external validation of the Bruker IR-BT® *Salmonella serogroup classifier v3*

Consec. No.	Strain No.	Serogroup Letter/Name	Serotype	Source
85	5362.2	*S.* D1 Lawndale	(1),9,12:z:1,5	European legless lizard, fecal sample
86	8782	*S.* D1 Lome	9,12:r:z6	Royal python
87	3238.2	*S.* D1 Panama	(1),9,12:l,v:1,5	Lizard
88	2553.2	*S.* D1 *ssp. enterica monophasic*	9,12:l,v:-	Pig, fecal sample
89	122.2	*S.* D1 *ssp. salamae*	9,12:z29:1,5	Turtle
90	30311.2	*S.* D1 Wangata	(1),9,12:z4,z23:-	Cow
91	33131	*S.* D1 Enteritidis	(1),9,12:g,m:-	Sushi roll
92	33274	*S.* D1 Enteritidis	(1),9,12:g,m:-	Falcon
93	34824	*S.* D1 Enteritidis	(1),9,12:g,m:-	Pig, fecal sample
94	34826	*S.* D1 Enteritidis	(1),9,12:g,m:-	Chicken, heart
95	35000	*S.* D1 Enteritidis	(1),9,12:g,m:-	Cat, swab
96	35070	*S.* D1 Enteritidis	(1),9,12:g,m:-	Laying hen, sock swab
97	35122	*S.* D1 Enteritidis	(1),9,12:g,m:-	Fanaloka, fecal sample
98	35135	*S.* D1 Enteritidis	(1),9,12:g,m:-	Polecat, small intestine
99	35176	*S.* D1 Enteritidis	(1),9,12:g,m:-	Hedgehog, lung
100	11141.2	*S.* D1 Enteritidis	(1),9,12:g,m:-	Alpaca
101	5560	*S.* D2 Baildon	9,46:a:e,n,x	Pacific boa
102	4815	*S.* D2 Quakam	9,46:z29:-	Sesame seed
103	35012	*S.* E1 Give	3,10:l,v:1,7	Dog, fecal sample
104	35073	*S.* E1 Muenster	3,10:e,h:1,5	Dog, swab
105	33893	*S.* E4 Senftenberg	1,3,19:g,s,t:-	Halva
106	9446	*S.* E4 Liverpool	1,3,19:d:e,n,z15	Hedgehog
107	30550	*S.* F Leeuwarden	11:b:1,5	Hedgehog, fecal sample
108	369.4	*S.* F *ssp. houtenae*	11:z4,z23,z32:	Iguana
109	33780	*S.* G Havana	(1),13,23:f,g:-	Pistachios
110	5270.3	*S.* G *ssp. enterica monophasic*	13,23:i:-	Poultry
111	11555.2	*S.* H Martonos	6,14,24:d:1,5	Laying hen, sock swab
112	4871.3	*S.* H *ssp. diarizonae*	14:z10:z	Milk snake
113	35077	*S.* I Hvittingfoss	16:b:e,n,x	Pasta dish
114	35625	*S.* I *ssp. salamae*	16:g,t:-	Bearded dragon, organ
115	32901	*S.* J *ssp. diarizonae*	17:l,v:z35	Snake, intestine
116	3505	*S.* J Jangwani	17:a:1,5	Bearded dragon, swab
117	35068	*S.* K Cerro	18:z4,z23:-	Dog, fecal sample
118	11193	*S.* K Sinthia	18:z38:-	Reptile
119	8148.2	*S.* L *ssp. arizonae*	21:z4,z23:-	Eyelash viper
120	5883.4	*S.* L Minnesota	21:b:e,n,x	Moringa capsules
121	732.2	*S.* M Mundonobo	28:d:1,7	Ceylon pit viper
122	6882	*S.* M Halle	28:c:1,7	Greek tortoise
123	1718	*S.* N Urbana	30:b:e,n,x	Western swamp tortoise
124	9334	*S.* N *ssp. salamae*	30:l,z28:z6	Skink
125	1473.2	*S.* O Adelaide	35:f,g:-	*Boa constrictor*
126	3881	*S.* O Adelaide	35:f,g:-	Bearded dragon

Table 1 (cont.). *Salmonella* (*S.*) strains used for the external validation of the Bruker IR-BT® *Salmonella serogroup classifier v3*

Consec. No.	Strain No.	Serogroup Letter/Name	Serotype	Source
127	4672	*S.* P Inverness	38:k:1,6	European legless lizard, fecal sample
128	7149	*S.* P *ssp. houtenae monophasic*	38:z4,z23:-	*Boa constrictor*
129	2189	*S.* R *ssp. houtenae*	40:z4,z32:-	snake
130	8418	*S.* R *ssp. salamae*	40:g,m,t:-	Anaconda
131	4907	*S.* S Offa	41:z38:-	Melon seeds
132	4976.3	*S.* S *ssp. arizonae*	41:z4,z23:-	Corn snake, organ
133	5529.3	*S. ssp. enterica, rough form*	-:-:	Pig, organ
134	32015	*S. ssp. houtenae, rough form*	-:-:-	Bearded dragon, peritoneum
135	5493	*S.* T *ssp. diarizonae*	42:k:z35	Corn snake
136	32483	*S.* T *ssp. salamae*	42:g,t:-	Peacock day gecko
137	32796	*S.* U *ssp. diarizonae*	43:z29:-	Rattle snake, abscess
138	456.3	*S.* U *ssp. houtenae*	43:z4,z23:-	Bearded dragon, organ
139	11433	*S.* V Guinea	(1),44:z10:1,7	Bearded dragon, fecal sample
140	32016	*S.* V *ssp. houtenae*	44:z4,z24:-	Bearded dragon, peritoneum
141	33139	*S.* W Apapa	45:m,t:-	Mexican west coast rattle-snake, fecal sample
142	10812.3	*S.* W *ssp. houtenae*	45:g,z51:-	Violaceous euphonia
143	8118.2	*S.* X *ssp. diarizonae*	47:i:z53	Kingsnake, organ
144	33326	*S.* X *ssp. salamae*	47:b:e,n,x,z15	Desert horned lizard
145	30643.2	*S.* Y *ssp. arizonae*	48:z4,z23:-	Racoon
146	4970.3	*S.* Y *ssp. diarizonae monophasic*	48:l,v:-	Kingsnake
147	6899.2	*S.* Z *ssp. arizonae*	50:z4,z23:-	Chain Kingsnake
148	30646.2	*S.* Z *ssp. diarizonae*	50:z:z52	Reptiles, fecal sample
149	4664	*S.* 51 *ssp. houtenae*	51:z4:z23	Boa
150	9887	*S.* 52 *ssp. diarizonae*	52:z:z52	Snake, organ
151	6261	*S.* 53 *ssp. arizonae*	53:z4,z23,z32:-	Bearded dragon
152	30491	*S.* 53 *ssp. diarizonae*	53:z10:z35	Nose-horned viper
153	3418	*S.* 55 *ssp. salamae*	55:k:z39	Leopard gecko
154	6615	*S.* 56 *ssp. arizonae*	56:z4,z23:-	Bearded dragon
155	8523	*S.* 56 *ssp. arizonae*	56:z4,z23,z32:-	Corn snake
156	884.2	*S.* 57 *ssp. diarizonae*	57:k:e,n,x,z15	Gecko
157	32739	*S.* 57 *ssp. diarizonae*	57:k:e,n,x,z15	Mountain kingsnake
158	35223	*S.* 58 *ssp. diarizonae*	58:z52:z35	Rock python, fecal sample
159	35403	*S.* 58 *ssp. salamae*	58:l,z13,z28:z6	Green basilisk, kidney
160	8304.2	*S.* 59 *ssp. diarizonae*	59:k:z	Gillen's dwarf monitor, liver
161	5598.2	*S.* 59 *ssp. diarizonae*	59:k:z	Horned viper, fecal sample
162	34995	*S.* 60 *ssp. diarizonae*	60:r:z	Bamboo snake, organ
163	35583	*S.* 60 *ssp. diarizonae*	60:Rz50:-	Kingsnake, kidneys
164	35494	*S.* 61 *ssp. diarizonae*	61:r:z53	Stimson's Python, fecal sample
165	35067	*S.* 61 *ssp. diarizonae, monophasic*	61:-:1,5,7	Sheep, organ
166	9888	*S.* 65 *ssp. diarizonae*	65:l,v:z	Snake, intestine
167	452.2	*S.* 65 *ssp. diarizonae*	65:z10:e,n,x,z15	*Boa constrictor*

Table 2. Proficiency test *Salmonella* (S.) strains used for challenging the Bruker IR-BT® *Salmonella serogroup classifier v3*

Consec. No.	Strain No.	Serogroup Letter/Name	Serotype	Year of Proficiency Test
1	1369	S. B Bredeney	4,12:i,v:1,7	2006
2	1394	S. B Typhimurium	(1),4,5,12:i:1,2	2006
3	10521	S. B Typhimurium	(1),4,5,12:i:1,2	2015
4	32798	S. B Typhimurium	(1),4,5,12:i:1,2	2020
5	33508	S. B Typhimurium	(1),4,5,12:i:1,2	2021
6	36162	S. B Typhimurium	(1),4,5,12:i:1,2	2024
7	8719.2	S. C1 Infantis	6,7:r:1,5	2024
8	30307	S. C1 Infantis	6,7:r:1,5	2017
9	35014	S. C1 Infantis	6,7:r:1,5	2023
10	2924	S. C1 Livingstone	6,7:d:l,w	2008
11	33269	S. C1 Oranienburg	6,7:m,t:-	2021
12	256.3	S. C2-C3 Kentucky	8,20:i:z6	2016
13	11723	S. C2-C3 Litchfield	6,8:l,v:1,2	2017
14	33540	S. C2-C3 Manchester	6,8:l,v:1,7	2021
15	5298	S. C2-C3 Manhattan	6,8:d:1,5	2009
16	10696.2	S. C2-C3 Newport	6,8:e,h:1,2	2016

tion of 6 cm^{-1}. It performed 32 scans for background and sample spectra. The zero filling factor was 1. This resulted in 1,814 data points in the aforementioned wave number range of 4,000 to 500 cm^{-1}. For post-processing, a "make compatible scalar" function was used for interpolating the spectrum, which resulted in a data set with exactly one data point per integer wavenumber, about 3,500 points in total.

After measurement, the resulting spectra were subjected to a quality test according to the instructions of the instrument manufacturer [33]. Spectra of poor quality with respect to e.g. minimum and maximum absorbance values and signal-noise-ratio were sorted out so that only spectra with acceptable quality were used for further analyses. Processing of spectra was performed using the IR Biotyper Client software (Bruker Daltonics) [33], applying version V4.0 and using the default settings recommended by the manufacturer, described as follows: The spectra were smoothed using the Savitzky-Golay algorithm over 9 data points and subsequently, the spectra's second derivative was calculated. Spectra were then cut to the relevant spectral window of 1,300 – 800 cm^{-1}. This window reflects the spectral absorbance region of various oligo- and polysaccharides [34], and hence the suitable spectral range to distinguish bacteria by their serogroups, i.e. outer membrane lipopolysaccharide-based structures. Finally, all spectra were vector-normalized in order to regulate preparation-related variance of biomass and hence absorption.

All qualitatively acceptable spectra were subjected to classification using the *Salmonella serogroup classifier v3* integrated into the IR Biotyper Client software version V4.0. According to the manufacturer, for this machine learning algorithm, a linear support vector machine (SVM) predictive model had been built. The predictive model was characterized by using the first 30 Principal Components (PC) and considering the spectral window of 1,300 – 800 cm^{-1} [33], which had been preselected as a wave number range suitable for discriminating between these serogroup classes (Norman Mauder, Bruker, personal communication). For the training of the classifier, n=158 *Salmonella* spp. isolates had been included, corresponding to n=36 O-serogroups.

The resulting serogroup classification of each spectrum was compared to the expected results, based on the serotyping of the *NRL-Salmonella*. In addition, the resulting numeric values for both the inlier and outlier scores [33], both intrinsic software parameters, were evaluated to assess the classification performance. The inlier and outlier scores indicate whether the recorded spectrum fits within the scope of spectral variance of the classifier, thus depicting the reliability of

the classification. If the outlier score is lower than the inlier threshold, the spectra are located in the spectral space of the training set. If the outlier score is between the inlier and outlier threshold, the spectra are located in the periphery of the spectral space of the training set. A classification score below the outlier threshold of 3.5 indicates a valid classification, while values >3.5 indicate that the sample spectrum is not covered by the spectral variance that was used to establish the classifier, rendering the sample spectrum out of scope and thus questionable.

2.4. Spectra exploration method 2D/3D scatter plots

In order to visualize the distribution of the salmonella serogroups in the IR spectral space, two- and three-dimensional scatter plots were generated using the IR Biotyper software V4.0 [33]. For this, the recorded single spectra were used in the case of the proficiency strain set (Table 2, p. 8). For the validation strain set (Table 1, p. 4), in order to decrease visual complexity, all acceptable-quality spectra of one respective isolate were assembled together into one average spectrum. With the respective single or average spectra, 2D and 3D scatter plots were created using the 2nd derivative of the spectra, considering the spectral window of 1,300 – 800 cm^{-1} and the dimensionality reduction algorithm PCA.

2.5. Evaluation of results

For quantitative evaluation of the results, the following parameters were noted for each sample spectrum, specified by the IR Biotyper software [33]: 1) the correctness of the serogroup classification in comparison to the actual serotype, and 2) the outlier score. A correct and valid classification was recorded if the attribution matched the actual serotype and the outlier score was ≤3.5. A questionable allocation was noted if the outlier score was >3.5, independent of the actual result assigned. An incorrect classification was observed if the allocation did not match the actual serotype provided that the outlier score was ≤3.5, respectively (supplementary material S1, S2, S3, S4).

The validation concept performed was inspired by the BVL "Guidelines for Validating Species Identifications using MALDI-ToF-MS" in a single laboratory [27, 28]. The evaluation of the external validation for the *Salmonella* serogrouping method using the IR Biotyper was performed as a so-called targeted identification. In order to assess the inclusivity, at least 20 strains of the target parameter needed to be validated with a true positive rate of ≥95 % and a false negative rate of ≤1 %. At the same time, the exclusivity was assessed by at least 30 validated strains of the non-target parameters achieving a true negative rate of ≥99 % and a false positive rate of ≤1 %. In this case, the target parameter was the respective serogroup under consideration, whereas the entirety of the other serogroups, including the Noah's Ark strain group, represented the non-target parameter. Thus, this evaluation was carried out for each desired target parameter separately within the study set under consideration (Figures 1 – 4, pp. 10 – 13). As an additional requirement, at least 90 % of both the target and the non-target parameter spectra had to be classified, i.e. a maximum of 10 % of questionable results was admissible.

3. Results

3.1. Classifier set-up

To build the *Salmonella serogroup classifier v3*, a support vector machine (SVM) predictive model was built by means of the IR Biotyper software 4.0. For the training of the classifier, *n*=158 *Salmonella* spp. isolates were included, corresponding to 36 O-serogroups and cultivated on seven different, widely used culture media. For the verification of the classifier, different internal (belonging to the study design) and external (independent) datasets were used [24].

3.2. Classifier validation

The *Salmonella serogroup classifier v3* was subsequently validated by a procedure inspired by the "Guidelines for Validating Species Identifications by MALDI ToF MS" [28], which was previously successfully applied for the validation of a *Listeria monocytogenes* serogroup classifier [29]. For this, a total of *n*=630 spectra from *n*=100 strains from four different target groups (Table 1) were classified using the above-mentioned classifier and all results duly noted (suppl mat S1, S2, S3, S4). The respective outgroup for each target parameter consisted of both the other three respective serogroups under consideration and an additional large group of *n*=409 spectra from a large set of *n*=67 strains from *n*=35 different non-targeted serogroups. Since two different serotype representatives were randomly chosen per serogroup, this group was termed "Noah's Ark strains". The Noah's Ark strains served as a sort of background noise for challenging the *Salmonella serogroup classifier v3* with a large set of diverse outgroup strains that might occur rarely, but nevertheless realistically in daily lab life.

Targeted Identification

following the guidelines for validating species identifications using MALDI-TOF-MS pursuant to §64 of the German Food and Feed code (LFGB), coordinated by the Federal Office of Consumer Protection and Food Safety, Berlin, Germany (BVL) 28 Oct 2022

Salmonella O-group B

Version:	1		
generated by	OBH	**checked by**	Dyk
generated on	8-Feb-2024	**checked on**	9-Feb-2024
Application	Identification of Salmonella O-groups		
Environment	Salmonella O:4 (B) and non - O:4 (B)		
Applied database / software version	Bruker classifier Salmonella O-groups v3 IR Biotyper v4.0 01/2023		

Validation set

number of target parameter isolates: 25	number of non-target parameter isolates: 142
number of target parameter spectra: 162	number of non-target parameter spectra: 877

cultivation conditions: COL-SB BD 37°C 24h aerobic

filtration method: preparation acc. to producer instructions (Instruction for use IR Biotyper Kit, Revision B (03/2021))

Only identified samples are used for determination of the inclusivity and exclusivity.

True-positive: outlier score </=3,5 AND target parameter classification = label	True negative: outlier score </=3,5 AND non-target parameter classification <> target parameter
False negative: true-positive criteria aren't fullfilled	False positive: true-negative criteria aren't fullfilled

Validation result	**Target parameter:**	***Salmonella O-group B***		target fulfilled	
Total number of spectra	162				
of which with ID result	**162**	*rate of identified spectra*	*100.0%*	90%	yes
Total number of isolates	25			20	yes
With respect to identified spectra:		**Inklusivity**			
Positive conformity (PA?)	162	**True positive rate**	**100.0%**	95%	yes
Negative deviation (ND)	0	**False negative rate**	**0.0%**	1%	yes
	Non-target parameter				
Total number of spectra	877				
of which with ID result	**840**	*rate of identified spectra*	*95.8%*	90%	yes
Total number of isolates	142			30	yes
With respect to identified spectra:		**Exklusivity**			
Positive conformity (PA?)	839	**True negative rate**	**99.9%**	99%	yes
Negative deviation (ND)	1	**False-Positive Rate**	**0.1%**	1%	yes

Out of 162 identified target parameter spectra, 162 (=100%) were classified correctly (inclusivity). 0 (=0%) of identified target parameter spectra were allocated incorrectly to a different serogroup.

Out of 840 identified non-target parameter spectra, 839 (=99.9%) were counted correctly as negative (exclusivity). 1 (=0.1%) of all spectra were misclassified as the target parameter.

Criteria for the **targeted identification** are thus **fulfilled**

Comment:

Release for site	**for department**	**on (date)**	**by (signature)**
CVUA S	T, MT, D	9-Feb-2024	RAU

The procedure is suitable for the intended purpose and is thereby released for usage.

Analytical parameter added to LIMS: ☐ yes ☑ no

A signed Version of this validation report is kept within the generating lab.

Figure 1. Summary report of the targeted identification of *Salmonella* serogroup B (O:4) following the Guidelines for Validating Species Identifications using MALDI-ToF-MS pursuant to §64 of the German Food and Feed code (LFGB), coordinated by the Federal Office of Consumer Protection and Food Safety, Berlin, Germany (BVL) 28 Oct 2022

Targeted Identification

following the guidelines for validating species identifications using MALDI-TOF-MS pursuant to §64 of the German Food and Feed code (LFGB), coordinated by the Federal Office of Consumer Protection and Food Safety, Berlin, Germany (BVL) 28 Oct 2022

Salmonella O-group C1

Version:	1		
generated by	OBH	**checked by**	Dyk
generated on	8-Feb-2024	**checked on**	9-Feb-2024
Application	Identification of Salmonella O-groups		
Environment	Salmonella O:7 (C1-C4) and non - O:7 (C1-C4)		
Applied database / software version	Bruker classifier Salmonella O-groups v3 IR Biotyper v4.0 01/2023		

Validation set

number of target parameter isolates: 25	number of non-target parameter isolates: 142
number of target parameter spectra: 150	number of non-target parameter spectra: 889

cultivation conditions: COL-SB BD 37°C 24h aerobic

filtration method: preparation acc. to producer instructions (Instruction for use IR Biotyper Kit, Revision B (03/2021))

Only identified samples are used for determination of the inclusivity and exclusivity.

True-positive: outlier score </=3,5 AND target parameter classification = label	True negative: outlier score </=3,5 AND non-target parameter classification <> target parameter
False negative: true-positive criteria aren't fullfilled	False positive: true-negative criteria aren't fullfilled

Validation result	**Target parameter:**	***Salmonella O-group C1***		target	fulfilled
Total number of spectra	150				
of which with ID result	**148**	*rate of identified spectra*	*98.7%*	90%	yes
Total number of isolates	25			20	yes
With respect to identified spectra:		**Inklusivity**			
Positive conformity (PA?)	147	**True positive rate**	**99.3%**	95%	yes
Negative deviation (ND)	1	**False negative rate**	**0.7%**	1%	yes
	Non-target parameter				
Total number of spectra	889				
of which with ID result	**854**	*rate of identified spectra*	*96.1%*	90%	yes
Total number of isolates	142			30	yes
With respect to identified spectra:		**Exklusivity**			
Positive conformity (PA?)	854	**True negative rate**	**100.0%**	99%	yes
Negative deviation (ND)	0	**False-Positive Rate**	**0.0%**	1%	yes

Out of 148 identified target parameter spectra, 147 (=99.3%) were classified correctly (inclusivity). 1 (=0.7%) of identified target parameter spectra were allocated incorrectly to a different serogroup.

Out of 854 identified non-target parameter spectra, 854 (=100%) were counted correctly as negative (exclusivity). 0 (=0%) of all non-target parameter spectra were misclassified as the target parameter.

Criteria for the **targeted identification** are thus **fulfilled**

Comment:

Release for site	**for department**	**on (date)**	**by (signature)**
CVUA S	T, MT, D	9-Feb-2024	RAU

The procedure is suitable for the intended purpose and is thereby released for usage.

Analytical parameter added to LIMS: ☐ yes ☑ no

A signed Version of this validation report is kept within the generating lab.

Figure 2. Summary report of the targeted identification of Salmonella serogroup C1 (O:7) following the Guidelines for Validating Species Identifications using MALDI-ToF-MS pursuant to §64 of the German Food and Feed code (LFGB), coordinated by the Federal Office of Consumer Protection and Food Safety, Berlin, Germany (BVL) 28 Oct 2022.

Targeted Identification

following the guidelines for validating species identifications using MALDI-TOF-MS pursuant to §64 of the German Food and Feed code (LFGB), coordinated by the Federal Office of Consumer Protection and Food Safety, Berlin, Germany (BVL) 28 Oct 2022

Salmonella O-group D1

Version:	1		
generated by	OBH	**checked by**	Dyk
generated on	8-Feb-2024	**checked on**	9-Feb-2024
Application	Identification of Salmonella O-groups		
Environment	Salmonella O:9 (D1) and non - O:9 (D1)		
Applied database / software version	Bruker classifier Salmonella O-groups v3 IR Biotyper v4.0 01/2023		

Validation set

number of target parameter isolates: 25	number of non-target parameter isolates 142
number of target parameter spectra: 159	number of non-target parameter spectra: 880

cultivation conditions: COL-SB BD 37°C 24h aerobic

filtration method: preparation acc. to producer instructions (Instruction for use IR Biotyper Kit, Revision B (03/2021))

Only identified samples are used for determination of the inclusivity and exclusivity.

True-positive: outlier score </=3,5 AND target parameter classification = label	True negative: outlier score </=3,5 AND non-target parameter classification <> target parameter
False negative: true-positive criteria aren't fullfilled	False positive: true-negative criteria aren't fullfilled

Validation result	**Target parameter:**	***Salmonella O-group D1***		target	fulfilled
Total number of spectra	159				
of which with ID result	**159**	*rate of identified spectra*	*100.0%*	90%	yes
Total number of isolates	25			20	yes
With respect to identified spectra:		**Inklusivity**			
Positive conformity (PA?)	159	**True positive rate**	**100.0%**	95%	yes
Negative deviation (ND)	0	**False negative rate**	**0.0%**	1%	yes
	Non-target parameter				
Total number of spectra	880				
of which with ID result	**843**	*rate of identified spectra*	*95.8%*	90%	yes
Total number of isolates	142			30	yes
With respect to identified spectra:		**Exklusivity**			
Positive conformity (PA?)	837	**True negative rate**	**99.3%**	99%	yes
Negative deviation (ND)	6	**False-Positive Rate**	**0.7%**	1%	yes

Out of 159 identified target parameter spectra, 159 (=100%) were classified correctly (inclusivity). 0 (=0%) of identified target parameter spectra were allocated incorrectly to a different serogroup.

Out of 843 identified non-target parameter spectra, 837 (=99.3%) were counted correctly as negative (exclusivity). 6 (=0.7%) of all non-target parameter spectra were misclassified as the target parameter.

Criteria for the **targeted identification** are thus **fulfilled**

Comment:

Release for site	**for department**	**on (date)**	**by (signature)**
CVUA S	T, MT, D	9-Feb-2024	RAU

The procedure is suitable for the intended purpose and is thereby released for usage.

Analytical parameter added to LIMS: ☐ yes ☑ no

A signed Version of this validation report is kept within the generating lab.

Figure 3. Summary report of the targeted identification of *Salmonella* serogroup D1 (O:9) following the Guidelines for Validating Species Identifications using MALDI-ToF-MS pursuant to §64 of the German Food and Feed code (LFGB), coordinated by the Federal Office of Consumer Protection and Food Safety, Berlin, Germany (BVL) 28 Oct 2022.

Targeted Identification

following the guidelines for validating species identifications using MALDI-TOF-MS pursuant to §64 of the German Food and Feed code (LFGB), coordinated by the Federal Office of Consumer Protection and Food Safety, Berlin, Germany (BVL) 28 Oct 2022

Salmonella O-group C2-C3

Version:	1		
generated by	OBH	**checked by**	Dyk
generated on	8-Feb-2024	**checked on**	9-Feb-2024
Application	Identification of Salmonella O-groups		
Environment	Salmonella O:8 (C2-C3) und non - O:8 (C2-C3)		
Applied database / software version	Bruker classifier Salmonella O-groups v3 IR Biotyper v4.0 01/2023		

Validation set

number of target parameter isolates: 25	number of non-target parameter isolates: 142
number of target parameter spectra: 159	number of non-target parameter spectra: 880

cultivation conditions: COL-SB BD 37°C 24h aerobic

filtration method: preparation acc. to producer instructions (Instruction for use IR Biotyper Kit, Revision B (03/2021))

Only identified samples are used for determination of the inclusivity and exclusivity.

True-positive: outlier score </=3,5 AND target parameter classification = label	True negative: outlier score </=3,5 AND non-target parameter classification <> target parameter
False negative: true-positive criteria aren't fullfilled	False positive: true-negative criteria aren't fullfilled

Validation result	**Target parameter:**	***Salmonella O-group C2-C3***		target	fulfilled
Total number of spectra	159				
of which with ID result	**154**	*rate of identified spectra*	*96.9%*	90%	yes
Total number of isolates	25			20	yes
With respect to identified spectra:		**Inklusivity**			
Positive conformity (PA?)	148	**True positive rate**	**96.1%**	95%	yes
Negative deviation (ND)	6	**False negative rate**	**3.9%**	1%	no
	Non-target parameter				
Total number of spectra	880				
of which with ID result	**848**	*rate of identified spectra*	*96.4%*	90%	yes
Total number of isolates	142			30	yes
With respect to identified spectra:		**Exklusivity**			
Positive conformity (PA?)	848	**True negative rate**	**100.0%**	99%	yes
Negative deviation (ND)	0	**False-Positive Rate**	**0.0%**	1%	yes

Out of 154 identified target parameter spectra, 148 (=96.1%) were classified correctly (inclusivity). 6 (=3.9%) of identified target parameter spectra were allocated incorrectly to a different serogroup.

Out of 848 identified non-target parameter spectra, 848 (=100%) were counted correctly as negative (exclusivity). 0 (=0%) of all non-target parameter spectra were misclassified as the target parameter.

Criteria for the **targeted identification** are thus **not fulfilled**

Comment:

Release for site	**for department**	**on (date)**	**by (signature)**
CVUA S	T, MT, D	9-Feb-2024	RAU

~~**The procedure is suitable for the intended purpose and is thereby released for usage.**~~

Analytical parameter added to LIMS: ☐ yes ☑ no

A signed Version of this validation report is kept within the generating lab.

Figure 4. Summary report of the targeted identification of *Salmonella* serogroup C2-C3 (O:8) following the Guidelines for Validating Species Identifications using MALDI-ToF-MS pursuant to §64 of the German Food and Feed code (LFGB), coordinated by the Federal Office of Consumer Protection and Food Safety, Berlin, Germany (BVL) 28 Oct 2022.

Concerning the target parameter *serogroup* O:4 (B) (Figure 1, p. 10), all *n*=162 spectra were indeed correctly identified as O:4. Out of *n*=877 non-target parameter spectra, *n*=840 (=95.8 %) were classified, with all but one spectrum being correctly identified as not belonging to serogroup O:4. Therefore, this parameter's sensitivity was 100 % and exclusivity was 99.9 %.

Similar results were noted for both serogroups O:6,7 (C1) and O:9 (D1):

With respect to *serogroup* O:6,7 (C1) (Figure 2, p. 11), *n*=148 out of *n*=150 target parameter spectra were correctly allocated to serogroup C1, corresponding to a true-positive rate of 99.3 %. Of the non-target parameter spectra, 96.1 % were identified and their classifications proven correct in each case, amounting to a true-negative rate of 100 %.

Regarding *serogroup* O:9 (D1) (Figure 3, p. 12), both the rate of identified target parameter spectra and their true-positive rate were 100 %. With respect to the non-target parameter, the exclusivity was 99.3 %, i.e. from *n*=843 identified spectra, all *n*=6 spectra from isolate CVUAS 30643.2 (*S. enterica* ssp. *arizonae* 48:z4,z23:-) were incorrectly assigned to the target parameter. Nevertheless, as the resulting false-positive rate of 0.7 % is below the threshold of 1 %, this target parameter *serogroup* O.9 (D1) was considered fully validated.

In this validation, the results of the target parameters *serogroup* O:4, *serogroup* O:6,7 and *serogroup* O:9 were fully in line with the adapted requirements of the national Guidelines [28]. These were used subsidiarily for validation and can therefore be considered as successfully validated by this challenging scheme (Figures 1 – 3, pp. 10 – 12).

With respect to the target parameter *serogroup* O:8 (C2-C3) (Figure 4, p. 13), while 96.1 % of the target parameter's spectra were correctly assigned, the classifier misallocated all *n*=6 spectra from strain CVUAS 4586 (*S.* Eboko 6,8:b:1,7) to different serogroups. This incorrect classification resulted in a false-negative rate of 3.9 %; thus, this parameter did not meet the requirements of the adapted Guidelines for Validating Species Identifications [28]. Regarding the non-target parameter spectra, all of the 96.4 % identified spectra were correctly classified as not belonging to serogroup O:8.

The false-negative identification of strain CVUAS 4586 suggests that the classifier in its current form does not cover the full biodiversity of this particular serogroup O:8. Therefore, the inclusion of more strains into the training set would most likely improve the classifier's performance and render it more powerful to accurately recognize this particular serogroup.

The quality of differentiation performance found in the current study has been observed before: Results for true positive rates were in a range previously described for *Salmonella* serovar differentiation by FTIR spectroscopy [15, 18, 20 – 22, 24 – 25]. Recently, Napoleoni *et al*. noted a sensitivity of 100 % for serogroup B, 91.5 % for serogroup C1 and 98.2 % for serogroup D1 [26].

With respect to the proficiency test strain set (Table 2, p. 8) used in the classifier trial, each of the spectra was correctly allocated to its respective serogroup.

3.3. Visualization of spectra in the spectral space

In the 2D scatter plot depicting all *n*=64 proficiency strain set single spectra (Table 2) from this study, the separation of the different classes is clearly visible (Figure 5) with respect to their most discriminatory Principal Component values PC 1 and 2. This demonstrates the capability of the trained classifier to differentiate between strains belonging to the specific serogroups under investigation.

The 3D scatter plot showing all *n*=167 validation strain set average spectra (Table 1, p. 4) in relation to their most prominent Principal Components PC1, PC2 and PC3 shows a more intricate picture (Figure 6). While the spectra clusters of serogroups O:9 and O:4 are comparatively large and well separated from the other spectra within the spectral space extending from PC 1 to PC 3, the spectra clusters of serogroups O:6,7 and O:8 are much more confined and thus depict less spectral variance. The spectra from serogroups O:6,7 and O:8 also intermingle more with the Noah's Ark strain spectra in this visualization. Nevertheless, as the external validation demonstrates, the AI-based classifier had been trained successfully to work out the discriminatory features of serogroup O:6,7 and, with one deviating strain also of serogroup O:8.

4. Concluding Remarks

The suitability of the *Salmonella serogroup classifier v3* for distinguishing between different *Salmonella* serogroups has thus been demonstrated for serogroups O:4, O:6,7 and O:9, three serogroups occurring very frequently in Europe overall and also in our daily lab practice. The fourth serogroup under investigation, O:8, failed to meet the adapted Guidelines [28] by only

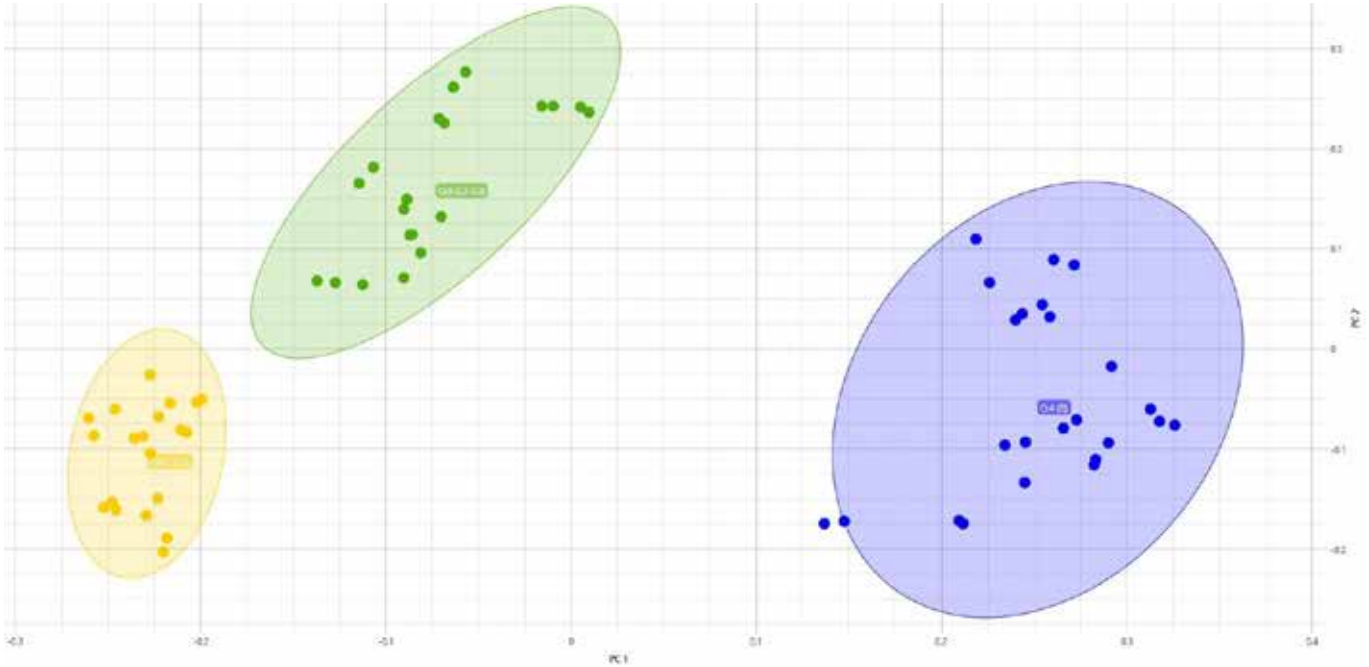


Figure 5. 2D-scatter plot of the proficiency test strains' single spectra (Table 2, p. 8), 2nd derivative, spectral range 1,300 – 800 cm^{-1}, dimensionality reduction algorithm PCA, displaying Principal Components PC 1 and PC 2. Blue: serogroup O4(B); green: serogroup =8(C2-C3); yellow: serogroup O6,7(C1)

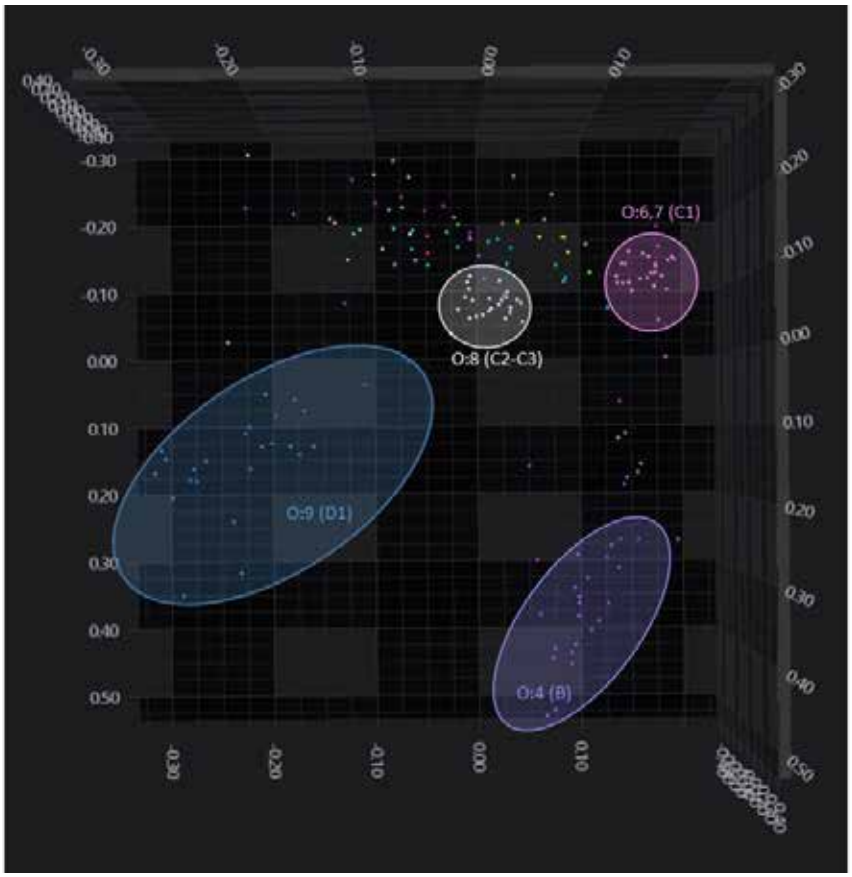


Figure 6. 3D-scatter plot of the validation strains' average spectra (Table 1, p. 4), 2nd derivative, spectral range 1,300 – 800 cm^{-1}, dimensionality reduction algorithm PCA, displaying Principal Components PC 1, PC 2 and PC 3. Colorful non-marked dots denote Noah's Ark strains.

one strain out of *n*=25 target parameter strains, suggesting good possibility of improvement. The classifier can now be used within our accredited lab workflow.

Applying the adapted Guidelines for the Validation of Species Identifications by MALDI ToF MS [28] provides users of any spectroscopy or spectrometric classification method with clear guidance for carrying out any external validations, thereby simplifying the implementation of new methods.

The Guidelines demand that large datasets be analyzed with respect to both the target parameter and the non-target parameter, thereby emulating conditions close to reality in an analytical laboratory with a high sample throughput. The classification method under investigation was thus tested and its performance evaluated under demanding conditions in order to obtain a robust and reliable result. Any classifier passing this performance test has therefore been proven to achieve the desired goal. Recently, a *Listeria* serogroup classifier has been proven by this method to reliably distinguish between different serogroups of *Listeria monocytogenes* [29].

For a future perspective, a serovar specific sub-classifier would enable the discrimination of selected serovars of the same serogroup. For instance, the specific determination of the very frequent Serovars *S.* Typhimurium (O:4), *S.* Infantis (O:6,7) and *S.* Enteritidis (O:9) would be both helpful in saving two days' time for analysis and in decreasing the amount of rabbit sera, both necessary for conventional serotyping. An analytical approach using FTIR spectroscopy that is faster than conventional serotyping would be an attractive method, especially for analyzing livestock with frequently repetitive isolates, such as e.g. *Salmonella* vaccination strains in chicken barns.

Acknowledgments

The authors wish to thank P. Bayan and L. Loeffler, both of CVUA Stuttgart, for their excellent technical assistance.

Statements

MC is a Bruker Daltonics employee. The other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. While Bruker Daltonics, Germany, set up the commercial classifier *Salmonella serogroup classifier v3*, Bruker Daltonics had no influence on the design of the validation nor on the collection, analyses or interpretation of the validation data. This study does not involve any human participants or animal experiments.

Data availability

Data will be made available on request. Metadata of the isolates used in the validation are listed in the MALDI-UP catalogue, https://maldi-up.ua-bw.de, including functional spreadsheet data files of all validations performed in the context of this study.

Supplementary material

Supplementary data associated with this article are available for this paper at https://doi.org/10.48414/aspects2025/16.

References

[1] European Centre for Disease Prevention and Control. Salmonellosis. Annual Epidemiological Report for 2022. 2024. Available from: https://www.ecdc.europa.eu/en/publications-data/salmonellosis-annual-epidemiological-report-2022 (last accessed 24 Sept 2024).

[2] U.S. Food & Drug Administration. Bad Bug Book. 2012. Available from: https://www.fda.gov/food/foodborne-pathogens/bad-bug-book-second-edition (last accessed 24 Sept 2024).

[3] German Federal Institute for Risk Assessment. The salmonella control programme is working: fewer salmonella detected on poultry. 2011. Available from: https://www.bfr.bund.de/en/press-release/the-salmonella-control-programme-is-working-fewer-salmonella-detected-on-poultry (last accessed 24 Sept 2024).

[4] European Food Safety Authority. Salmonella. Available from: https://www.efsa.europa.eu/en/topics/topic/salmonella#:~:text=In%20the%20EU%2C%20over%2091%2C000,it%20can%20be%20life%2Dthreatening. (last accessed 24 Sept 2024).

[5] Bertrand S., Rimhanen-Finne R., Weill F. X., Rabsch W., Thornton L., Perevoscikovs J., van Pelt W., Heck M., Salmonella infections associated with reptiles: the current situation in Europe. Euro Surveill, 2008. 13(24), https://www.ncbi.nlm.nih.gov/pubmed/18761944

[6] EFSA Panel on Biological Hazards, Scientific Opinion on the risk posed by Shiga toxin-producing Escherichia coli (STEC) and other pathogenic bacteria in seeds and sprouted seeds. EFSA Journal, 2011. 9(11): p. 2424 DOI: https://doi.org/10.2903/j.efsa.2011.2424.

[7] Grimont P.A.D. & Weill F.-X. Antigenic Formulae of the Salmonella Serovars. 2007. Available from: https://www.pasteur.fr/sites/default/files/veng_0.pdf).

[8] Issenhuth-Jeanjean S., Roggentin P., Mikoleit M., Guibourdenche M., de Pinna E., Nair S., Fields P. I., Weill F. X., Supplement 2008–2010 (no. 48) to the White-Kauffmann-Le Minor scheme. Res Microbiol, 2014. 165(7): p. 526–30 DOI: https://doi.org/10.1016/j.resmic.2014.07.004.

[9] Ferrari R.G., Rosario D.K.A., Cunha-Neto A., Mano S.B., Figueiredo E.E.S., Conte-Junior C.A., Worldwide epidemiology of Salmonella serovars in animal-based foods:

a meta-analysis. Applied and environmental microbiology, 2019. 85(14): p. e00591 – 19,
[10] Wenning M., Scherer S., Identification of microorganisms by FTIR spectroscopy: perspectives and limitations of the method. Appl Microbiol Biotechnol, 2013. 97(16): p. 7111 – 20 DOI: https://doi.org/10.1007/s00253-013-5087-3.
[11] Quintelas C., Ferreira E. C., Lopes J. A., Sousa C., An Overview of the Evolution of Infrared Spectroscopy Applied to Bacterial Typing. Biotechnol J, 2018. 13(1) DOI: https://doi.org/10.1002/biot.201700449.
[12] Novais A., Freitas A. R., Rodrigues C., Peixe L., Fourier transform infrared spectroscopy: unlocking fundamentals and prospects for bacterial strain typing. Eur J Clin Microbiol Infect Dis, 2019. 38(3): p. 427 – 448 DOI: https://doi.org/10.1007/s10096-018-3431-3.
[13] Santos C. Fraga M.E., Kozakiewicz Z., Lima, N., Fourier transform infrared as a powerful technique for the identification and characterization of filamentous fungi and yeasts. Research in Microbiology, 2010. 161(2): p. 168 – 175 DOI: https://doi.org/10.1016/j.resmic.2009.12.007.
[14] Seltmann G., Voigt W., Beer W., Application of physico-chemical typing methods for the epidemiological analysis of Salmonella enteritidis strains of phage type 25/17. Epidemiology & Infection, 1994. 113(3): p. 411 – 424,
[15] Kim S., Reuhs B. L., Mauer L. J., Use of Fourier transform infrared spectra of crude bacterial lipopolysaccharides and chemometrics for differentiation of Salmonella enterica serotypes. J Appl Microbiol, 2005. 99(2): p. 411 – 7 DOI: https://10.1111/j.1365-2672.2005.02621.x.
[16] Baldauf N.A. Rodriguez-Romo L.A., Yousef, A.E., Rodriguez-Saona, L.E., Differentiation of Selected Salmonella enterica Serovars by Fourier Tranform Mid-Infrared Specktroscopy. Appl Spectroscopy, 2006. 60(6): p. 592 – 598,
[17] Baldauf N.A. Rodriguez-Romo L.A., Männig, A., Yousef, A.E., Rodriguez-Saona, L.E., Effect of selective growth media on the differentiation of Salmonella enterica serovars by Fourier-Transform Mid-Infrared Spectroscopy. J Microbiol Methods, 2007. 68: p. 106 – 114,
[18] Männig A. Baldauf N.A., Rodriguez-Romo L.A., Yousef, A.E., Rodriguez-Saona L.E., Differentiation of Salmonella enterica Serovars and Strains in Cultures and Food Using Infrared Spectroscopic and Microspectroscopic Techniques Combined with Soft Independent Modeling of Class Analogy Pattern Recognition Analysis. J Food Protection, 2008. 71(11): p. 2249 – 2256,
[19] De Lamo-Castelvi S.D. Männig A., Rodiguez-Saona L.E., Fourier-tranform infrared sepctroscopy combined with ummunomagnetic separation as a tool to discriminate Salmonella serovars. Analyst, 2010. 135: p. 2987 – 2992,
[20] Preisner O., Guiomar R., Machado J., Menezes J. C., Lopes J. A., Application of Fourier transform infrared spectroscopy and chemometrics for differentiation of Salmonella enterica serovar Enteritidis phage types. Appl Environ Microbiol, 2010. 76(11): p. 3538 – 44 DOI: https://doi.org/10.1128/AEM.01589-09.
[21] Mauder N., Rau J., Identification of Salmonella enterica serovar Enteritidis vaccine strains by Fourier transform infrared spectroscopy. Aspects of Food Control and Animal Health, 2014. 1: p. 2 – 14,
[22] Campos J., Sousa C., Mourao J., Lopes J., Antunes P., Peixe L., Discrimination of non-typhoid Salmonella serogroups and serotypes by Fourier Transform Infrared Spectroscopy: A comprehensive analysis. Int J Food Microbiol, 2018. 285: p. 34 – 41 DOI: https://doi.org/10.1016/j.ijfoodmicro.2018.07.005.
[23] Oberreuter H., Rau J., Artificial neural network-assisted Fourier transform infrared spectroscopy for differentiation of Salmonella serogroups and its application on epidemiological tracing of Salmonella Bovismorbificans outbreak isolates from fresh sprouts. FEMS Microbiol Lett, 2019. 366(15) DOI: https://doi.org/10.1093/femsle/fnz193.
[24] Cordovana M., Mauder N., Join-Lambert O., Gravey F., LeHello S., Auzou M., Pitti M., Zoppi S., Buhl M., Steinmann J., Frickmann H., Dekker D., Funashima Y., Nagasawa Z., Soki J., Orosz L., Veloo A. C., Justesen U. S., Holt H. M., Liberatore A., Ambretti S., Pongolini S., Soliani L., Wille A., Rojak S., Hagen R. M., May J., Pranada A. B., Kostrzewa M., Machine learning-based typing of Salmonella enterica O-serogroups by the Fourier-Transform Infrared (FTIR) Spectroscopy-based IR Biotyper system. J Microbiol Methods, 2022. 201: p. 106564 DOI: https://doi.org/10.1016/j.mimet.2022.106564.
[25] Cordovana M., Mauder N., Kostrzewa M., Wille A., Rojak S., Hagen R. M., Ambretti S., Pongolini S., Soliani L., Justesen U. S., Holt H. M., Join-Lambert O., Hello S. L., Auzou M., Veloo A. C., May J., Frickmann H., Dekker D., Classification of Salmonella enterica of the (Para-)Typhoid Fever Group by Fourier-Transform Infrared (FTIR) Spectroscopy. Microorganisms, 2021. 9(4) DOI: https://doi.org/10.3390/microorganisms9040853.
[26] Napoleoni M., Ceschia S., Mitri E., Beneitez E. E., Silenzi V., Staffolani M., Rocchegiani E., Blasi G., Gurian E., Identification of Salmonella Serogroups and Distinction Between Typhoidal and Non-Typhoidal Salmonella Based on ATR-FTIR Spectroscopy. Microorganisms, 2024. 12(11) DOI: https://doi.org/10.3390/microorganisms12112318.
[27] Rau J., Eisenberg T., Wind C., Huber I., Pavlovic M., Becker R., From the § 64 LFGB MALDI-TOF working group: guidelines for validation of species identifications with MALDI-TOF-MS. J. Consum. Prot. Saf., 2021. 17(1): p. 97 – 101 DOI: https://doi.org/10.1007/s00003-021-01353-x.
[28] Rau J., Dolch L.J., Eisenberg T., Erhard M., Fuchs J., Gödecke P., Hilgarth M., Huber I., Huschek G., Neumann N., Mailänder M., Pavlovic M., Stahl A., Wind C., Wittmann C., Becker R.; Guidelines for validating species identifications using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) in a single laboratory or in laboratory networks, coordinated by the Federal Office of Consumer Protection and Food Safety MALDI-TOF-MS working group pursuant to § 64 of the German Food and Feed Code (LFGB), Berlin (BVL), Editor. 28 Oct 2022.

[29] Oberreuter H., Dyk M., Rau J., Validated differentiation of Listeria monocytogenes serogroups by FTIR spectroscopy using an Artificial Neural Network based classifier in an accredited official food control laboratory. Clinical Spectroscopy, 2023. 5 DOI: https://doi.org/10.1016/j.clispe.2023.100030.
[30] Microbiology of the food chain — Horizontal method for the detection, enumeration and serotyping of Salmonella — Part 1: Detection of *Salmonella* spp. [ISO 6579-1:2017]. 2017. Available from: https://www.iso.org/standard/56712.html).
[31] LRQA Nederland B.V., MALDI Biotyper® complete solution for the confirmation of Salmonella spp.. 12 Feb 2018. Available from: https://assets.bettyblocks.com/2e826c3765f4401c84671a06de961fc7_7ec-88453307c4ce7af6f9258f9645792/21479/2017 LR73_-_MALDI_Biotyper_Salmonella_spp._-_Certificate__2022-2026_.pdf (last accessed 11 March 2025).
[32] Bruker Daltonics, Instructions for Use - IR Biotyper Kit. REF 1851760, Mar 2021, (Doc. no. 5023357 Revision B)
[33] Bruker Daltonics, IR Biotyper Software User Manual. REF 1845471, May 2022, (Doc. no. 5025119 Revision F)
[34] Naumann D., Infrared Spectroscopy in Microbiology, in Encyclopedia of Analytical Chemistry, R.A. Meyers, Editor. 2000, John Wiley & Sons Ltd: Chichester. p.102–131.