Food Microbiology 91 (2020) 103534

Contents lists available at ScienceDirect

Food Microbiology

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

# A genoserotyping system for a fast and objective identification of *Salmonella* serotypes commonly isolated from poultry and pork food sectors in Belgium

Mathieu Gand[a,b], Wesley Mattheus[a,*], Nancy Roosens[c], Katelijne Dierick[d], Kathleen Marchal[b,e], Sophie Bertrand[f,1], Sigrid C.J. De Keersmaecker[c,1]

[a] Sciensano, Infectious Diseases in Humans, Bacterial Diseases, B-1180, Brussels, Belgium
[b] Department of Information Technology, IDLab, imec, Ghent University, B-9052, Ghent, Belgium
[c] Sciensano, Transversal Activities in Applied Genomics, B-1050, Brussels, Belgium
[d] Sciensano, Infectious Diseases in Humans, Food Pathogen, B-1050, Brussels, Belgium
[e] Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052, Ghent, Belgium
[f] Sciensano, B-1050, Brussels, Belgium

ARTICLE INFO

*Keywords:*
*Salmonella*
Genoserotyping
Luminex
Decision support system
Food safety

ABSTRACT

Humans are mostly contaminated by *Salmonella* through the consumption of pork- and poultry-derived food products. Therefore, a strict monitoring of *Salmonella* serotypes in food-producing animals is needed to limit the transmission of the pathogen to humans. Additionally, *Salmonella* can lead to economic loss in the food sector. Previously, a genoserotyping method using the MOL-PCR and Luminex technology was developed for the identification of the 6 *Salmonella* serotypes, and their variants, subjected to an official control in the Belgian food sector. In this study, 3 additional assays using the same technology were developed for the rapid and cost-effective detection of 13 dangerous highly invasive serotypes or other serotypes frequently isolated from the Belgian poultry and pork sector, i.e. Agona, Anatum, Brandenburg, Choleraesuis, Derby, Enteritidis vaccine strains, Gallinarum var. Gallinarum/Pullorum, Livingstone, Mbandaka, Minnesota, Ohio, Rissen and Senftenberg. Moreover, the previously developed first MOL-PCR assay was improved for *S.* Paratyphi B and serogroup O:3 detection. Finally, a Decision Support System hosted by a web application was created for an automatic and objective interpretation of the Luminex raw data. The 3 new assays and the modifications of the first assay were validated with a 100% accuracy, using 553 *Salmonella* and non-*Salmonella* strains in total.

## 1. Introduction

In 2018 salmonellosis was once again considered as the second foodborne disease in Europe (EFSA, 2019). Its causing agent, *Salmonella*, can infect a large variety of food-producing animals like poultry, pigs and cattle, which are their major reservoirs. While the transmission between animals is oro-fecal, humans are infected by contact with animals and mostly by the consumption of contaminated food (estimated to account for 85% in transmissions to humans) (Heredia and García, 2018; Oxford Analytica, 2012). The *Salmonella* genus is divided into more than 2500 serotypes (following the Kaufman-White-Le Minor (KWL) scheme) among which the 1500 of the *Salmonella enterica* subsp. *enterica* are generally responsible for food poisoning (Antunes et al., 2016; Grimont and Weill, 2007; Ryan et al., 2017). The *Salmonella* serotypes have different host-specificities. Depending on the *Salmonella* serotype and the host infected by it, the salmonellosis can result either in non-invasive symptoms like gastroenteritis (or even silent symptoms for healthy carriers) or in more dangerous invasive symptoms like fever and bacteremia, leading potentially to the decease of the host without treatment (Heredia and García, 2018). Indeed, while *Salmonella* serovar Enteritidis is more associated with poultry, *Salmonella* serovar Typhimurium can infect a broader range of animal species including poultry, pork and cattle. In poultry, these 2 non-invasive serotypes will lead to low or undetectable symptoms, allowing the bacteria to infect humans through the consumption of contaminated eggs and meat (Demirbilek, 2016). Contrarily, *Salmonella* serovar Gallinarum biovar Gallinarum (*S.* Gallinarum var. Gallinarum) and *Salmonella* serovar Gallinarum biovar Pullorum (*S.* Gallinarum var. Pullorum) are restricted to some avian species (depending on their age) and cause invasive symptoms like severe septicemia, resulting in a high mortality rate in a.o. the poultry

* Corresponding author. Sciensano, rue Engeland 642, 1180, Brussels.
*E-mail address:* wesley.mattheus@sciensano.be (W. Mattheus).
[1] Equally contributed.

https://doi.org/10.1016/j.fm.2020.103534
Received 23 January 2020; Accepted 22 April 2020
Available online 27 April 2020
0740-0020/ © 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

sector (Alves Batista et al., 2018). Similarly, *Salmonella* serovar Choleraesuis is historically known to cause large outbreaks especially in pigs, leading to septicemia, enterocolitis and pneumonia. Despite the fact that these invasive *Salmonella* serotypes are rare in Europe, their surveillance must be maintained to detect eventual reemergence, like the outbreak caused by *S.* Choleraesuis in 4 Danish pig farms in 2012–2013 (Pedersen et al., 2015). *S.* Enteritidis and *S.* Typhimurium are the most prevalent serotypes in developed countries, although infections by *S.* Enteritidis decreased these last years, thanks to the vaccination obligation of poultry breeding with live attenuated vaccines (Griffin and O'Brien, 2013 + NRC personnal communication) such as Salmovac SE (IDT Biologika, Dessau, Germany) or AviPro SALMONELLA VAC E (Elanco GmbH, Cuxhaven, Germany).

While the presence of non-invasive *Salmonella* serotypes must be monitored in food-producing animals to limit transmissions to human, the rapid detection of invasive serotypes is more needed to avoid economical loss for the breeders, linked to the decrease in productivity (weight loss, abortions, milk production, treatment of contaminated eggs, etc…), the quarantining of diseased animals and the destruction or treatment of contaminated food products (Heredia and García, 2018; Majowicz et al., 2010; Oxford Analytica, 2012). For example, in poultry farms, the farmers must check the presence of *Salmonella* in the flock before sending the animals to the slaughterhouse. The Turn-Around Time (TAT), which means the maximum period allowed to communicate *Salmonella* serotyping results by the first line laboratories, after collection of the samples at the farm, is established by the Belgian Federal Agency for the Security of the Food Chain (FASFC) at 14 days for poultry breeding. This time includes the sampling, the transport to the first line laboratories, the isolation of the *Salmonella* and the serotype identification for which a referring to the National Reference Center (NRC) is generally needed. During this period, the animal transfer to the slaughterhouse or the selling of the eggs is postponed until a positive agreement is obtained from the first line laboratories. If a *Salmonella* is isolated from the field, corrective actions must be undertaken depending on the serotype identified, going from logistic slaughtering to restriction of eggs selling to applications that involve a thermic treatment. A complete disinfection of the farm must be performed in each case and absence of *Salmonella* must be proven on site before the arrival of a new animal batch.

Consequently, it is crucial to rapidly identify the serotype of the *Salmonella* isolated from poultry and pork sectors, with the aim to reduce human food poisoning, but also to quickly react in case of outbreaks due to invasive serotypes such as Choleraesuis and Gallinarum. The gold standard method for *Salmonella* serotyping consists of the characterization of 3 antigenic sites (somatic antigen O and flagellar antigens H1 and H2), located at the surface of the bacterium, by slide agglutination with specific antisera. More than 120 antisera are needed to be able to identify all of the 2500 serotypes included in the KWL scheme (Grimont and Weill, 2007). A positive agglutination is not always clearly obtained, making the result interpretation subjective. Additionally, for the discrimination of some specific variants (e.g. *Salmonella* serovar Paratyphi B variant Java, *S.* Gallinarum var. Gallinarum and *S.* Gallinarum var. Pullorum), biochemical tests based on culturing methods are needed. Also, when isolating *S.* Enteritidis, it can be important to make the discrimination between the vaccine and the wild type field strain, if the vaccination campaign was too close to the sampling period. The vaccine strains included in the vaccines AviPro SALMONELLA VAC E and Salmovac SE are respectively resistant to antibiotics (streptomycin and rifampicin) and auxotrophic double-mutant (ade- and his-). Therefore, their differentiation is made by testing their growth characteristics on specific media containing antibiotics or lacking adenine and histidine. But as these biochemical and growth tests are based on culturing on specific media, they are complex, time-consuming and not always reliable (Batista et al., 2013; Gand et al., 2019; Maurischat et al., 2015). This is why all these reference methods (i.e. slide-agglutination, biochemical and growth tests) are only fully mastered in the NRCs, which dispose of the totality of the antisera collection, and where the tests are performed by experienced and carefully trained technicians. For this reason, most of the *Salmonella* isolates must be sent to the NRC for a complete identification and this causes additional delays not always compatible with the short TAT asked by the regulation. Therefore, this situation is not suitable for the animal sectors. This is why there is a need for a faster, cheaper and more accurate identification technique, which could be used by the NRCs to reduce their TAT and analysis costs, or directly by the first line laboratories, avoiding the need to send the sample to another laboratory.

Luckily, alternative appropriate methods exist for the identification of *Salmonella* serotypes and their variants, among the new molecular tools developed these last years. Some target-based molecular methods, such as Multiplex Oligonucleotide Ligation – PCR (MOL-PCR) linked to the Luminex technology, have proven to be suitable and cost-effective for rapid diagnostics (Jean-Gilles Beaubrun et al., 2014; Liang et al., 2016; Yoshida et al., 2016a). The MOL-PCR allows the detection of genomic molecular markers, linked to the serotype, by specific probes through a ligation-amplification reaction, at a high multiplexing level. The so created MOL-PCR fragments are then hybridized to unique color-coded MagPlex beads, subsequently detected by a MagPix apparatus, based on a fluorescence reaction (Luminex xTAG technology). An assay using this technology was previously developed by Gand et al. (2020) for a fast, objective and cost-effective genoserotyping of 6 *Salmonella* serotypes (and their variants) mentioned in the Belgian regulation (Belgian royal decree 27/04/2007 and Belgian FASFC note BP-FDS/LABO/1470050 v7) and for the determination of common serogroups (O:3,10, O:4, O;7, O:8, O:9 and O:21).

In the present study, we have developed 3 new MOL-PCR assays with the aim to (i) rapidly detect highly invasive *Salmonella* serotypes like *S.* Gallinarum var. Gallinarum, *S.* Gallinarum var. Pullorum and *S.* Choleraesuis, (ii) make the discrimination between wild type and vaccine (AviPro SALMONELLA VAC E and Salmovac SE) isolates of *S.* Enteritidis and (iii) identify common serotypes isolated in the food chain including: *Salmonella* serovar Agona, *Salmonella* serovar Anatum, *Salmonella* serovar Brandenburg, *Salmonella* serovar Derby, *Salmonella* serovar Livingstone, *Salmonella* serovar Mbandaka, *Salmonella* serovar Minnesota, *Salmonella* serovar Ohio, *Salmonella* serovar Rissen and *Salmonella* serovar Senftenberg. Additionally, the first MOL-PCR assay described by Gand et al. (2020) was improved for a more specific detection of *S.* Paratyphi B var. Java and serogroup O:3. The molecular markers, specific to the serotypes targeted by the developed assays, were selected from the MultiLocus Sequence Typing (MLST) database (EnteroBase), from the scientific literature or based on in-house genomic comparison studies using Whole Genome Sequencing (WGS). In addition, a Decision Support System (DSS) hosted by a web-application was created. The aim of this DSS is to perform an automatic interpretation of the Luminex raw data and to give recommendations to the users in case of partial identification. Through this web-application, all the final identification results are also centralized in a database for national surveillance of the *Salmonella* serotypes circulating in Belgium. The *Salmonella* genoserotyping system, including the multiplex assays and the DSS, was compared to the classical methods (slide-agglutination, growth and biochemical tests) with the analysis of 553 *Salmonella* and non-*Salmonella* strains.

## 2. Mat & meth

### 2.1. Bacterial strains and DNA preparation

Seventeen isolates of *S.* Choleraesuis and 2 isolates of *S.* Gallinarum were respectively provided by the Belgian Institute of Tropical Medicine (Antwerp, Belgium) and the company Biovac (Beaucouzé, France). The serotype of these isolates was confirmed by the NRC prior to use. All the other *Salmonella* strains used in this study have

previously been isolated from food, animal or human samples (between 2005 and 2018 in Belgium) and were sent to the NRC for further characterization (including serotype identification), after *Salmonella* genus confirmation by the first line laboratories. These isolates, in addition to the non-*Salmonella* strains used in this study, were stored in the collection of the NRC and are available upon request (Supplementary Table S4). All isolates were cultured on Nutrient agar (Neogen® Culture Media, Lansing, USA).

For MOL-PCR, the bacterial DNA of samples and controls were extracted by heat lysis as described by Gand et al. (2020). For WGS, genomic DNA was extracted with the GenElute Bacterial Genomic DNA kit (Sigma-Aldrich, Saint-Louis, USA) according to the manufacturer's instructions.

### 2.2. Selection of molecular markers from WGS data

Using an Illumina MiSeq instrument (2 × 300 bp, Nextera XT libraries), genomic DNA of 11 *S.* Livingstone, 1 *S.* Gallinarum var. Gallinarum and 4 *S.* Gallinarum var. Pullorum isolates was sequenced. The FASTQ reads were deposited at the SALMSTID BioProject on NCBI (PRJNA509747).

In CLC Genomics Workbench 8.0 (Qiagen, Hilden, Germany), the raw FASTQ reads were first trimmed to a quality score limit of 0.05 with maximum 2 ambiguous nucleotides and reads with a length below 30 nucleotides were discarded. These trimmed reads were then *de novo* assembled with automatic bubble and word size, in mapping mode "map reads back to contigs" with scaffolding and a minimum contig length of 1000 nucleotides. Identically, the downloaded Sequence Reads Archive (SRA) (Supplementary Table S1) were trimmed and assembled as described for the in-house sequenced data. All assemblies were uploaded to the *Salmonella In Silico* Typing Resource (SISTR), developed by Yoshida et al. (2016b), for serotype confirmation and MLST typing.

The downloaded and in-house produced WGS data were all exported to Gegenees (version 2.2.1; downloaded from http://www.gegenees.org; Ågren et al., 2012) on a Linux platform and analyzed for the selection of molecular markers specific of targeted serotypes, as described by Gand et al. (2020). All the genomes were labelled in the software as target, reference or background as indicated in the Supplementary Table S1.

### 2.3. Selection of molecular markers from enterobase and scientific literature

Single Nucleotide Polymorphism (SNP) markers were selected among the allele sequences of the 7 housekeeping genes (*aroC*, *dnaN*, *hemD*, *hisD*, *purE*, *sucA*, and *thrA*) of the MLST database EnteroBase (Achtman et al., 2012; Alikhan et al., 2018; https://enterobase.warwick.ac.uk), as previously described by Gand et al. (2020) (Table 1). The genetic alignments were made using MUSCLE (Edgar, 2004) in the MEGA7 software (Kumar et al., 2016). The specificity of the alleles conserved in close Sequence Types (STs) or eBurst Groups (eBGs) was checked *in silico* on an in-house curated version of Enterobase downloaded in early 2019 and composed of 186 900 entries at that time. Abs/Pres markers (based on the presence or absence of a genetic sequence) and other SNP markers, presented in Table 2, were inspired from genetic studies or molecular methods published in the scientific literature.

### 2.4. Ligation probe design and MOL-PCR protocol

The ligation probes presented in Tables 1 and 2 were designed using the Visual OMP (version 7.6.58.0; DNA Software) according to the guidelines of Wuyts et al. (2015). When the molecular markers were selected from the literature, the ligation probes were designed based on existing probes, primers or a specific amplified sequence, if available. These probes were ordered from Integrated DNA Technologies (IDT, Leuven, Belgium) with a standard desalted purification.

The MOL-PCR reactions, the hybridization to MagPlex-TAG microspheres (Luminex, Austin, USA), the staining reaction using streptavidin-R-phycoerythrin (SAPE) (Thermofischer Scientific, Waltham, USA) and the read-out using a Mag-Pix device (Luminex, Austin, USA) were performed following the protocol detailed by Gand et al. (2020). For the 4 MOL-PCR assays, the negative control (CTRL_-) was composed of *Vibrio alginolyticus* DNA (strain M/5035) extracted and used identically as for the other samples. For positive controls composed of DNA belonging to several *Salmonella* serotypes, one colony per serotype was mixed in a tube at the DNA extraction step (section 2.1) and processed like the other samples. The positive controls (CTRL_+) used for each assay, and their composition, are listed in Supplementary Table S2.

### 2.5. Data interpretation using the DSS

The MagPix produces Median Fluorescence Intensity (MFI) for each probe and each sample. From these raw data, the Signal to Noise Ratios (SNRs) are calculated for all the molecular markers (Abs/Pres and SNP markers). When it is needed, an Allele Call (AC) is also performed for some SNP markers. Then, these processed data are converted into a serotype identification based on a barcode system: the Gödel Prime Product (GPP) (Van den Bulcke et al., 2008; Van Den Bulcke et al., 2010). The data interpretation, including the checking of quality controls (CTRL_- and CTRL_+), the processing of raw data, the conversion into serotype identification and the serogroup clustering (using the GPP), was previously described in detail by Gand et al. (2020). In this study, these operations were automatically performed by the DSS developed as a web-application and accessible (through a login request) at: https://salmstid.wiv-isp.be/. The web-application was developed and deployed according to the DTAP principle, i.e. following the 4 phases of Development, Testing, Acceptance and Production. The web-application was first configured for each MOL-PCR assay, using a protocol setting function (included in it) in which all the cut-off values, the quality controls limits, the serogrouping probes and the correspondences between GPPs and identification results can be set. Then, the comma separated value (.csv) file generated by the MagPix for each assay, containing all the MFI results for each probe and sample, was uploaded in the DSS. After the automatic processing of the data, the results were displayed on the screen and were exported as a PDF report. The performance of the DSS was tested with many scenarios simulating different case studies which can happen during routine analyses such as normal identifications, serogroup clustering, mixed (not pure) sample, failed quality controls, fluorescence detection issues or incorrect setting of the MagPix apparatus.

Briefly, all the serotype identifications are linked to a GPP, itself linked to a specific combination of positive probes which include at least the detection of the marker *invA* (for the detection of *Salmonella* spp.) and a marker targeting one of the following serogroups: O:3, O:4, O:7, O:8, O:9, O:21. When obtaining a GPP configured in the DSS, the system displays the name of the detected serotype. If the marker *invA* is not retrieved, the sample is reported as "No *Salmonella*" by the DSS. If it is present in the sample but without a serogroup marker, the *Salmonella* isolate is characterized as "Unknown serogroup". On the opposite, if more than one serogroup marker is detected (which is not possible considering the serogroup targeted by the method) in a *Salmonella* sample (positive for *invA*), the DSS will report it as "not pure sample" because a mix of serotypes is suspected. In case of a sample with an unknown GPP and not belonging to one of the cases described above, the DSS will display the serogroup to which it belongs, based on the division of the GPP by the prime number of the probe, and recommend what further analysis must be performed to complete the identification. When the interpretation of the data cannot be properly done, because of failed quality controls, detection errors or wrong settings of the MagPix, the DSS displays an error message to the user, describing the anomaly (Fig. 1).

**Table 1**
Ligation probes designed from MLST markers selected in EnteroBase.

| Target | MOL-PCR assay | Probe | DNA sequence | Beads region[a] | Interpretation[b] | MLST gene | Allele number | SNP position[c] |
|---|---|---|---|---|---|---|---|---|
| *S.* Agona | O3-4-21 | STID20-U_SNP | *TAATACGACTCACTATAGG*gttgtaaattgtagtaaagaagtaCGGGCAAACGCGCGCTA | 15 | SNR | *hemD* | 7 | 249 |
| | | STID20-D | P-ATTTTGCGTGGCAATGGCG*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| | | STID21-U_SNP | *TAATACGACTCACTATAGGG*aataagagaattgatatgaagatgGCAGAAATACGGCCTGCGT | 35 | SNR | *aroC* | 3 | 53 |
| | | STID21-D | P-GATTACCGTGGCGGTGGACG*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| *S.* Anatum | O3-4-21 | STID22-U_SNP | *TAATACGACTCACTATAGGG*tgaaatgtgtatttgtatgtttagAAGTGGTTTCCGCTCATCGT | 62 | SNR | *purE* | 25 | 87 |
| | | STID22-D | P-ACCCCCGATAAGCTGTTCAGCTT*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| | | STID23-U_SNP | *TAATACGACTCACTATAGGG*agtagaaagttgaaattgattatgCTGAAAGAGATGGTTCGA | 12 | SNR | *sucA* | 20 | 132 |
| | | STID23-D | P-CATGCGGGTAACAGCGGCAC*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| *S.* Brandenburg | O7 | STID36-U_SNP* | *TAATACGACTCACTATAGG*gtaattgaattgaaagataagtgtCTTCCGCGCCATCGTAATGTTTT | 18 | SNR | hemD | 13 | 345 |
| | | STID36-D* | P-GCACATCGTTGATAACATTCACAAAAAC*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| | | STID37-U_SNP | *TAATACGACTCACTATAGG*gattgatatttgaatgtttgtttgCGGGCGGCGCGGCACAT | 22 | SNR | purE | 10 | 174 |
| | | STID37-D | P-CTGCCGGGAATGATTGCGGC*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| S. Choleraesuis | O7 | STID45-U_SNP | *TAATACGACTCACTATAGG*gttagttatgatgaatattgtgtaCCTGCGCGATTACCGTGGT | 45 | SNR | aroC | 36 | 66 |
| | | STID45-D | P-GGTGGACGTTCTTCCGCGC*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| S. Derby ST682 | O3-4-21 | STID48-U_SNP | *TAATACGACTCACTATAGG*gtaagagtattgaaattagtaagaCTGTCTACGCTGCCTGCCGT | 66 | SNR | dnaN | 60 | 191 |
| | | STID48-D | P-CGATTTCCCGAATCTTGACGAC*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| S. Mbandaka | O7 | STID10-U_SNP | *TAATACGACTCACTATAGG*gtgattgaatagtagattgtttaaCAACATGGCGATGTTTAGCGTT | 46 | SNR | thrA | 68 | 360 |
| | | STID10-D | P-TCCGGCCCAGGAATGAAAGGGATGAT*CCCTTTAGTGAGGGTTAAT* | | | | | |
| S. Mbandaka S. Rissen | O7 | STID11-U_SNP | *TAATACGACTCACTATAGGG*gtgttatagaagttaaatgttaagATCGGTAAAGCCGGTGCGGCTAAT | 30 | SNR | purE | 64 | 333 |
| | | STID11-D | P-GCCGCCCTGCTCGCCGCGCA*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| S. Minnesota | O3-4-21 | STID24-U_SNP | *TAATACGACTCACTATAGG*gtatgttgtaatgtattaagaaagAGATTGCCGTTCAGTTGGAAGGT | 25 | AC | dnaN | 11 | 132 |
| | | STID24-U_WT | *TAATACGACTCACTATAGGG*tattagagtttgagaataagtagtGAGATTGCCGTTCAGTTGGAAGGC | 33 | | | | |
| | | STID24-D | P-GATCGGATGCTGGTGCGTTCTGG*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| | | STID25-U_SNP | *TAATACGACTCACTATAGG*gtgattgaatagtagattgtttaaCTGCCGCGCGCGGACACT | 46 | SNR | hisD | 102 | 441 |
| | | STID25-D | P-GCCCGGCAGGCCCTGAGCGC*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| S. Ohio | O7 | STID46-U_SNP | *TAATACGACTCACTATAGG*gtaagattagaagttaatgaagaaCGCTGCGCGGCAGCCAGG | 52 | SNR | aroC | 82 | 475 |
| | | STID46-D | P-ATCGCGATGAAATCACGGCG*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| | | STID9-U_SNP* | *TAATACGACTCACTATAGGG*gtaattgaattgaaagataagtgtTACAAAGGCGTTGCCGGGA | 18 | SNR | hisD | 12 | 174 |
| | | STID9-D* | P-CCAAAAATTTTATCCACTTTCGGTACGGAC*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| S. Rissen | O7 | STID12-U_SNP | *TAATACGACTCACTATAGGG*tattagagtttgagaataagtagtTCGCCATGCGGGGTAACAGT | 33 | SNR | sucA | 151 | 147 |
| | | STID12-D | P-GGCACTCGCGAAGTGGTGCT*CCCTTTAGTGAGGGTTAAT* | | | | | |
| S. Senftenberg | O3-4-21 | STID7-U_SNP | *TAATACGACTCACTATAGGG*attgtgaaagaaagagaagaaattCCGCCATGTCTCGCGCCGGGATCTCA | 14 | SNR | thrA | 19 | 435 |
| | | STID7-D | P-GTGGTGCTCATTACCCAGTCCTCCTC*CCTTTAGTGAGGGTTAAT* | | | | | |
| | | STID26-U_SNP | *TAATACGACTCACTATAGG*aattagaagtaagtagagtttaagAAACGGAAGGTAGCGAACTGCGT | 56 | SNR | dnaN | 6 | 354 |
| | | STID26-D | P-ACTGTCGCGACCGACGGCCA*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| | | STID28-U_SNP | *TAATACGACTCACTATAGG*atagatttagaatgaattaagtgATCCTGCTGACGCCTGATGCTC | 28 | SNR | hisD | 75 | 376 |
| | | STID28-D | P-ACATTGCCCGCAAGGTGGCG*TCCCTTTAGTGAGGGTTAAT* | | | | | |
| | | STID29-U_SNP | *TAATACGACTCACTATAGGG*gtgttatagaagttaaatgttaagCGTCCAGCGACGATGATAACCTA | 30 | SNR | thrA | 64 | 312 |
| | | STID29-D | P-CCAGTTAAAGGGATCTCTAACCTTAAC*TCCCTTTAGTGAGGGTTAAT* | | | | | |

P-: Phosphate.
Primer (T7 and T3), anti-TAG, target-specific sequences and SNP positions are indicated by italic, lower-case, underlined and bold sequences, respectively.
*: the probes anneal on the reverse complement strand of the MLST gene.
[a] Corresponds to the specific color and TAG sequence of the bead.
[b] The presence of the molecular marker is determined by the calculation of the Signal to Noise Ratio (SNR) or Allele Call (AC).
[c] Corresponds to the location of the SNP in multiple alignments of all alleles related to the considered MLST gene.

**Table 2**
Ligation probes designed from molecular markers inspired from the scientific literature and WGS study.

| Target | MOL-PCR assay | Probe | DNA sequence | Beads region[a] | Type[b] | Interpretation[c] | Marker | Based on | Source |
|---|---|---|---|---|---|---|---|---|---|
| AviPro SALMONELLA VAC E | O9 | STID401-U | *TAATACGACTCACTATAGGGttt*gttagaatgagaagatttatgCGCCGCCAAAGACCATTTCGT | 75 | SNP | SNR | kdpA | kdpA-V_probe | Maurischat et al., 2015 |
| | | STID401-D | P-CAATTTGCATCAGCCACATCGGC*TCCCTTTAGTGAGGGTTAAT* | | | | | | |
| S. Brandenburg | O7 | STID38-U_SNP | *TAATACGACTCACTATAGG*gtaagattagaagttaatgaagaa GGGTGTTGGTGGTGCAACAAA | 52 | SNP | SNR | fliC | Genbank accession no. AY434709 | Herrera-León et al., 2004 |
| | | STID38-D | P-CGGTACTGCTGCAATTAAGGA*TCCCTTTAGTGAGGGTTAAT* | | | | | | |
| S. Derby | O3-4-21 | STID1-U | *TAATACGACTCACTATAGG*gttagttatgatgaatattgtgta TTGGAGATCTTTCTAATGCGGAT | 45 | Abs/Pres | SNR | troN | Genes of SPI-23 conserved in S. Derby | Hayward et al., 2013 |
| | | STID1-D | P-TCTACTAATACTGTCATCATGTTTGGAC*TCCCTTTAGTGAGGGTTAAT* | | | | | | |
| S. Enteritidis | O9* | STID2-U | *TAATACGACTCACTATAGGG*ttattagagtttgagaataagtagtCGGCGCATTCCTCCGTTT | 33 | Abs/Pres | SNR | sdr | Ligation probes STID2-U and STID2-D | Gand et al., 2020 |
| | | STID2-D | P-TTTCGTCGTGGGCGTCAGTA*TCCCTTTAGTGAGGGTTAAT* | | | | | | |
| S. Gallinarum | O9 | STID41-U_SNP | *TAATACGACTCACTATAGG*gatagatttagaatgaattaagtgCATTTATGGCCCCGGAGGCGA | 28 | SNP | SNR | hypothetical protein | Genomic study using Gegenees | This study |
| | | STID41-D | P-ATGACGGGGACGACACCTGGT*CCCTTTAGTGAGGGTTAAT* | | | | | | |
| S. Gallinarum var. Gallinarum | O9 | STID43-U_SNP | *TAATACGACTCACTATAGG*gtatgttgtaatgtattaagaaag CCGACTCTTCGCTGAATATTATGA | 25 | SNP | SNR | MscS | Genomic study using Gegenees | This study |
| | | STID43-D | P-TGTATTGCTTTACCAAAACAACAGTATCCCTTTAGTGAGGGTTAAT | | | | | | |
| S. Livingstone | O7 | STID47-U | *TAATACGACTCACTATAGG*atagatttagaatgaattaagtgCTGAGAACCTGATAAATCG | 28 | Abs/Pres | SNR | hypothetical protein | Genomic study using Gegenees | This study |
| | | STID47-D | P-CGGCTATCTGGAGCGCGT*CCCTTTAGTGAGGGTTAAT* | | | | | | |
| S. Paratyphi B (including var. Java) | BASE | STID333-U | *TAATACGACTCACTATAGGG*tgatatagtagtgaagaaataagt AGGACGTATCTCTTTAGCAGGC | 34 | Abs/Pres | SNR | SPAB_01124 | Reverse primer pPB23 | Zhai et al., 2014 |
| | | STID333-D | P-GGAGGGGAGAAAGATATTTATGCCT*CCCTTTAGTGAGGGTTAAT* | | | | | | |
| Salmonella | All | invA-U | *TAATACGACTCACTATAGGG*gataagaaagtgaaatgtaaattg ATAAACTTCATCGCACCGTCA | 51 | Abs/Pres | SNR | invA | Ligation probes invA-U and invA-D | Wuyts et al., 2015 |
| | | invA-D | P-AAGGAACCGTAAAGCTGGCTTT*CCCTTTAGTGAGGGTTAAT* | | | | | | |
| Salmovac SE | O9 | STID40-U_SNP | *TAATACGACTCACTATAGG*gttgtaaattgtagtaaagaagtaCGGTCTGACCTCCATGCTA | 15 | SNP | SNR | nhaA | nhaA_V_probe | Maurischat et al., 2015 |
| | | STID40-D | P-CCGTTGGGAATTATTGCCG*TCCCTTTAGTGAGGGTTAAT* | | | | | | |
| Serogroup O:4 | O3-4-21* | STID16-U | *TAATACGACTCACTATAGGG*tttgatttaagagtgttgaatgtaTCAAGTTGGAACTGGTGCT | 26 | Abs/Pres | SNR | rfbJ | Ligation probes STID16-U and STID16-D | Gand et al., 2020 |
| | | STID16-D | P-GGGGTAAGTTTGAAAGATTTTCTGG*TCCCTTTAGTGAGGGTTAAT* | | | | | | |
| Serogroup O:9 | O9* | STID171-U_SNP | *TAATACGACTCACTATAGGG*aattgagaaagagataaatgatag CATATACTAAACAAAAAGCAAATGAAC | 72 | SNP | SNR | prt | Ligation probes STID171-U and STID171-D | Gand et al., 2020 |
| | | STID171-D | P-TCGCCGCCGCCATTATAGA*TCCCTTTAGTGAGGGTTAAT* | | | | | | |
| Serogroup O:6,7 | O7* | STID18-U | *TAATACGACTCACTATAGG*gtaagagtattgaaattagtaaga CGTTGGCAGACTGGTACTGATTG | 66 | Abs/Pres | SNR | wbaA | Ligation probes STID18-U and STID18-D | Gand et al., 2020 |
| | | STID18-D | P-GCTCCCCTATTACGATGATTTC*TCCCTTTAGTGAGGGTTAAT* | | | | | | |
| Serogroup O:3 | O3-4-21* | STID301-U | *TAATACGACTCACTATAGGG*tgatatagtagtgaagaaataagt TCTCTACGCAGACAATTATGTCA | 39 | Abs/Pres | SNR | wzx | Primer E_wzx_F | Franklin et al., 2011 |
| | | STID301-D | P-TGGAGTTATTATCCGGATGGG*TCCCTTTAGTGAGGGTTAAT* | | | | | | |

(*continued on next page*)

**Table 2** (*continued*)

| Target | MOL-PCR assay | Probe | DNA sequence | Beads region[a] | Type[b] | Interpretation[c] | Marker | Based on | Source |
|---|---|---|---|---|---|---|---|---|---|
| Serogroup O:3,10 | O3-4-21 | STID31-U | *TAATACGACTCACTATAGG*Gaaataagaatagagagaagaaagtt TTATAAATTTAGGTTTAGAACATGTTTAC | 34 | Abs/Pres | SNR | wzx | O:3,10 sequence not present in O:1,3,19 | Zhang et al., 2015 |
| | | STID31-D | P-GGTGAGAGGATAAAGCAGGTAAAA*TCCCTTTAGTGAGGGTTAAT* | | | | | | |
| Serogroup O:1,3,19 | O3-4-21 | STID321-U | *TAATACGACTCACTATAGGG*gatatagtagtgaagaataagt TCTCTACGCAGACAATTATGTCA | 13 | Abs/Pres | SNR | wzx | O:1,3,19 sequence not present in O:3,10 | Zhang et al., 2015 |
| | | STID321-D | P-TGGAGTTATTATCCGGATGGG*TCCCTTTAGTGAGGGTTAAT* | | | | | | |
| Serogroup O:21 | O3-4-21* | STID35-U | *TAATACGACTCACTATAGG*gttgagaattagaatttgataaag CCACTGTCATTGGTGGTTATCAG | 73 | Abs/Pres | SNR | wzx | Ligation probes STID21-U and STID21-D | Gand et al., 2020 |
| | | STID35-D | P-TATGAATGGCTGGTATACGACATC*TCCCTTTAGTGAGGGTTAAT* | | | | | | |

P-: Phosphate.
Primer (T7 and T3), anti-TAG, target-specific sequences and SNP positions are indicated by italic, lower-case, underlined and bold sequences, respectively.
*: Additionally present in the module BASE (Gand et al., 2020).
[a] Corresponds to the specific color and TAG sequence of the bead.
[b] The presence of the molecular marker is determined by the calculation of the Signal to Noise Ratio (SNR) or Allele Call (AC).
[c] Corresponds to the location of the SNP in multiple alignments of all alleles related to the considered MLST gene.

### 2.6. Evaluation of the MOL-PCR assays

A total of 464 *Salmonella* isolates and 33 non-*Salmonella* isolates were used to evaluate the sensitivity (using inclusivity tests), the specificity (using exclusivity tests) and the accuracy of the 3 new developed MOL-PCR assays following the approach used by Gand et al. (2020). From the 464 *Salmonella* isolates, 330 belonged to the serotypes targeted by the method and were used for the inclusivity tests (Table 3). The remaining 134 *Salmonella* isolates, belonging to 75 other serotypes not targeted by the method (Tables 4 and 5), and 33 non-*Salmonella* isolates (Table 5), were used for the exclusivity tests.

A part of these isolates were also used for the validation of the modifications made to the MOL-PCR previously developed by Gand et al. (2020). This included all the isolates belonging to serogroup O:3 (56 to O:3,10 and 32 to O:1,3,19) and those belonging to serotype Paratyphi B (11) (Tables 3 and 4). In addition, 56 other *S.* Paratyphi B were also tested for the evaluation of these modifications (Table 6).

In case of discordances in the results obtained between the developed and the reference methods, 2 additional serotyping analyses by slide-agglutination were performed (one blind and one by another technician) in order to verify the reference identification.

## 3. Results

### 3.1. Selection of the molecular markers

For each *Salmonella* serotype, including its variant, specific SNP markers were screened in the housekeeping gene alleles of the MLST scheme (Achtman et al., 2012), which are conserved in a serotype population (Table 1). After having checked *in silico* the specificity of the candidates in EnteroBase, it appeared that a combination of at least 2 MLST markers (in association with the serogroup markers, Table 2 and Supplementary Table S3) was needed to reach a false positive rate close to 0% and a false negative rate lower than 10%. These values were later experimentally verified during the validation of the method (see section 3.3).

Additional molecular markers were selected or inspired from scientific papers describing a target-based genoserotyping method or a genomic comparison study (Table 2). The specificity of these markers was checked *in silico* by BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi) using the NCBI database.

As no appropriate markers could be selected in EnteroBase nor in the scientific literature for the specific detection of *S.* Livingstone and *S.* Gallinarum var. Gallinarum, genomic comparison studies were performed using publicly available and in-house generated WGS data (Supplementary Table S1). Fragmented alignments were performed using the software Gegenees for the determination of genomic signatures specific for each of the 3 target groups (*S.* Livingstone, *S.* Gallinarum and *S.* Gallinarum var. Gallinarum genomes) and absent in the background group (*Salmonella* genomes belonging to 43 other serotypes). For each of the 2 serotypes or variant, one molecular marker located in a coding sequence, suitable for a ligation probe design and offering a good specificity after a BLAST check in NCBI, was selected (Table 2).

### 3.2. Development of the MOL-PCR assays

The MOL-PCR is based on the principle that upstream and downstream probes must anneal close to each other on the target sequence to be subsequently linked by the ligase enzyme. For each of the previously selected molecular markers (in section 3.1), these probes were designed. Their sequence is listed in Tables 1 and 2. A wild type version of the upstream probe was also designed when an interpretation using AC was needed to improve the fluorescence detection. The probes developed in this study were divided in 3 different MOL-PCR modules, i.e. MOL-PCR O3-4-21, MOL-PCR O7 and MOL-PCR O9. In each module,

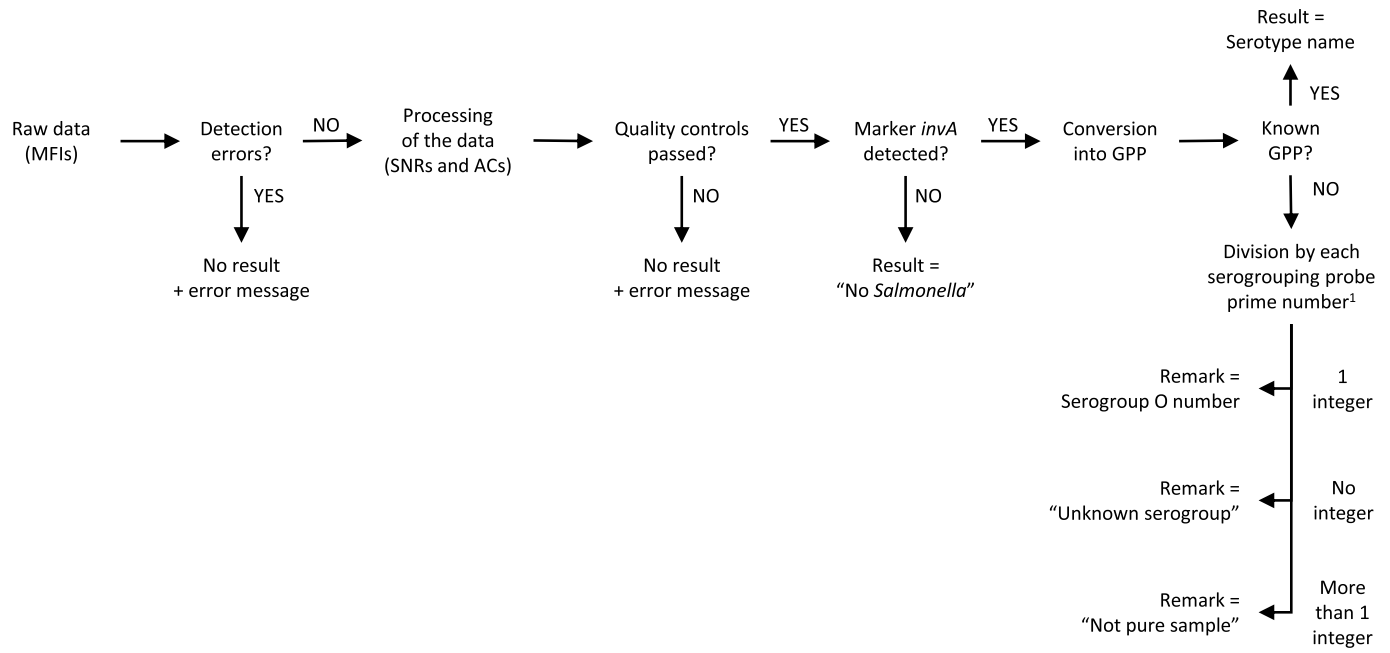


**Fig. 1.** Decision tree for results interpretation performed by the Decision Support System (DSS). AC: Allele Call; GPP: Gödel Prime Product; MFI: Median Fluorescence Intensity; SNR: Signal to Noise Ratio. From the upload of the comma separated value (.csv) file on the web application, the DSS performs the automatic interpretation of the raw data, using the GPP, to display identification results. The processing of the data includes experimental error detection and quality control checks (see section 2.5). [1]:to know if a specific molecular marker was detected through the probe combination resulting into a GPP, this latter is divided by the prime number of the probe targeting this molecular marker. If an integer is obtained, the molecular marker was detected. If a decimal number is obtained the molecular marker was not present.

serogrouping probes used to cluster each *Salmonella* isolate (positive for the marker *invA*) into one of the serogroups targeted by the modules, when possible, were included: STID16 (O:4), STID18 (O:7), STID35 (O:21), STID171 (O:9) and STID301 (O:3) (Table 2). With exception of the latter, all other serogrouping probes were already previously described in the context of the MOL-PCR assay developed by Gand et al. (2020), named here MOL-PCR BASE, allowing the detection of 6 *Salmonella* serotypes (and their variants) mentioned in the Belgian regulation, i.e. *S.* Enteritidis, *S.* Hadar, *S.* Infantis, *S.* Paratyphi B including the Java variant, *S.* Typhimurium including its monophasic variant and *S.* Virchow.

The module MOL-PCR O3-4-21 was developed to detect the *Salmonella* serotypes: Agona, Anatum, Brandenburg, Derby, Minnesota and Senftenberg. The probe STID48 was included in the assay for the identification of the particular sequence type ST682 of *S.* Derby (representing 10% of the Derby population in EnteroBase) not detected by STID1. Also in this assay, the detection of *S.* Senftenberg was obtained though the combination of STID7 and STID26, representing 86% of the Senftenberg population in Enterobase clustered in eBG55 and eBG69, or STID28 and STID29 representing 10% of the Senftenberg population in Enterobase clustered in eBG30. The probes STID31 and STID321 included in the MOL-PCR O3-4-21 were used for the discrimination between O:3,10 and O:1,3,19 respectively, but only when the *Salmonella* sample is already serogrouped as O:3 by STID301 (Supplementary Table S3a).

The module MOL-PCR O7 was created for the specific detection of the invasive *S.* Choleraesuis and other common *Salmonella* serotypes belonging to serogroup O:7: Livingstone, Mbandaka, Ohio and Rissen. The probe combination of invA (*Salmonella*), STID18 (O:7), STID9 and STID46 is used for the specific detection of *S.* Ohio. For the detection of *S.* Livingstone, the probes invA (*Salmonella*), STID18 (O:7) and STID47 must be positive. In the module O7, it can be noticed that the probe STID47 was also sometimes positive for *S.* Ohio but always together with STID9 and STID46 (Supplementary Table S3b).

The module MOL-PCR O9 was developed for a fast identification of the serotype Gallinarum, the discrimination between its 2 variants Gallinarum and Pullorum but also for the differentiation between the *S.* Enteritidis wild type field and vaccine strains. In this module, STID40 and STID401 are respectively used for the detection of Salmovac SE and AviPro SALMONELLA VAC E when the *S.* Enteritidis identification has been confirmed by the probes invA, STID2 and STID171 also included in the module O9. The discrimination between the variants Gallinarum and Pullorum is performed in this assay by STID43 and STID42, respectively, when the sample is already positive for invA, STID41 and STID171 (Supplementary Table S3c).

For all modules, the expected combinations of molecular markers, the GPPs, the associated serotyping results and the corresponding remarks are listed in the Supplementary Table S3. These parameters were set in the DSS for the automatic interpretation of the data.

### *3.3. Validation of the 3 new MOL-PCR assays by comparison with the classical method*

For the validation of the 3 new MOL-PCR assays (MOL-PCR O3-4-21, MOL-PCR O7 and MOL-PCR O9) and the DSS, at least 25 targeted *Salmonella* isolates (when available), 75 untargeted *Salmonella* isolates and 25 non-*Salmonella* isolates were analyzed per modules for the inclusivity and exclusivity tests. The identification results produced with the new method were compared with those obtained using the reference techniques (Table 3–5).

Using the MOL-PCR O3-4-21, 27 *S.* Agona, 28 *S.* Anatum, 25 *S.* Brandenburg, 46 *S.* Derby, 28 *S.* Minnesota and 28 *S.* Senftenberg were correctly completely identified by the DSS. A part of the *S.* Derby isolates (16, representing 34%) were reported as belonging to the sequence type ST682 which is more than the proportion of this cluster estimated previously in EnteroBase (10%). Identically for the *S.* Senftenberg detection, more isolates (25%) belonging to the eBG30 were identified compared to the percentage of this population in EnteroBase (10%). With the MOL-PCR O7, the complete identification of 17 *S.* Choleraesuis, 30 *S.* Livingstone, 29 *S.* Mbandaka, 26 *S.* Ohio

**Table 3**
Results of the inclusivity tests using *Salmonella* isolates belonging to the targeted serotypes.

| MOL-PCR assay | Number of tested isolates | Reference identification[a] | Serogroup | Expected identifications with the MOL-PCR assays | | | Obtained identifications with the MOL-PCR assays | | | Comparison expected vs. obtained |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | GPP | Serotype | Serogroup | GPP | Serotype | Serogroup | |
| O3-4-21 | 27 | *S.* Agona | O:4 | 2040753 | *S.* Agona | O:4 | 2040753 | *S.* Agona | O:4 | 27 TP |
| | 28 | *S.* Anatum | O:3,10 | 33915 | *S.* Anatum | O:3 | 33915 | *S.* Anatum | O:3 | 28 TP |
| | 25 | *S.* Brandenburg | O:4 | 35547369 | *S.* Brandenburg | O:4 | 35547369 | *S.* Brandenburg | O:4 | 25 TP |
| | 30 | *S.* Derby | O:4 | 31317 or 779493 | *S.* Derby | O:4 | 31317 | *S.* Derby | O:4 | 30 TP |
| | 16 | *S.* Derby | O:4 | 31317 or 779493 | *S.* Derby | O:4 | 779493 | *S.* Derby ST682 | O:4 | 16 TP |
| | 28 | *S.* Minnesota | O:21 | 572241 | *S.* Minnesota | O:21 | 572241 | *S.* Minnesota | O:21 | 28 TP |
| | 21 | *S.* Senftenberg | O:1,3,19 | 110055 or 189255 | *S.* Senftenberg | O:3 | 110055 | *S.* Senftenberg eBG55 eBG69 | O:3 | 21 TP |
| | 7 | *S.* Senftenberg | O:1,3,19 | 110055 or 189255 | *S.* Senftenberg | O:3 | 189255 | *S.* Senftenberg eBG30 | O:3 | 7 TP |
| O7 | 17 | *S.* Choleraesuis | O:7 | 435 | *S.* Choleraesuis | O:7 | 435 | *S.* Choleraesuis | O:7 | 17 TP |
| | 30 | *S.* Livingstone | O:7 | 1155 | *S.* Livingstone | O:7 | 1155 | *S.* Livingstone | O:7 | 30 TP |
| | 29 | *S.* Mbandaka | O:7 | 4845 | *S.* Mbandaka | O:7 | 4845 | *S.* Mbandaka | O:7 | 29 TP |
| | 18 | *S.* Ohio | O:7 | 2415 or 26565 | *S.* Ohio | O:7 | 26565 | *S.* Ohio | O:7 | 18 TP |
| | 8 | *S.* Ohio | O:7 | 2415 or 26565 | *S.* Ohio | O:7 | 2415 | *S.* Ohio | O:7 | 8 TP |
| | 27 | *S.* Rissen | O:7 | 3315 | *S.* Rissen | O:7 | 3315 | *S.* Rissen | O:7 | 27 TP |
| O9 | 1 | AviPro SALMONELLA VAC E | O:9 | 5865 | AviPro SALMONELLA VAC E | O:9 | 5865 | AviPro SALMONELLA VAC E | O:9 | 1 TP |
| | 5 | Enteritidis | O:9 | 255 | *S.* Enteritidis wild type | O:9 | 255 | *S.* Enteritidis wild type | O:9 | 5 TP |
| | 2 | *S.* Gallinarum var Gallinarum | O:9 | 1365 | *S.* Gallinarum var Gallinarum | O:9 | 1365 | *S.* Gallinarum var Gallinarum | O:9 | 2 TP |
| | 10 | *S.* Gallinarum var Pullorum | O:9 | 1155 | *S.* Gallinarum var Pullorum | O:9 | 1155 | *S.* Gallinarum var Pullorum | O:9 | 10 TP |
| | 1 | Salmovac SE | O:9 | 4845 | Salmovac SE | O:9 | 4845 | Salmovac SE | O:9 | 1 TP |
| TOTAL: | 330 | | | | | | | Total TP: | | 330 |
| | | | | | | | | Total FN: | | 0 |
| | | | | | | | | Inclusivity (sensitivity): | | 100% |

TP: True Positive; FN: False Negative.
[a] Obtained by classical methods, i.e. slide-agglutination, cultural and biochemical tests.

**Table 4**
Results of the exclusivity tests using untargeted *Salmonella* isolates for which the serogroup can be determined by the developed method.

| Number of tested isolates | Reference identification[a] | Serogroup | Results obtained with the MOL-PCR assays | | | | Comparison with the reference identification | |
|---|---|---|---|---|---|---|---|---|
| | | | Using the module[b] | GPP | Identification | Serogroup | Serotype identification | Serogroup clustering |
| 5 | *S.* 1,4,[5],12:i:- | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Abony | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Agama | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Bareilly | O:7 | O7 | 105 | *Salmonella* | O:7 | TN | OK |
| 1 | *S.* Berta | O:9 | O9 | 15 | *Salmonella* | O:9 | TN | OK |
| 1 | *S.* Braenderup | O:7 | O7 | 15 | *Salmonella* | O:7 | TN | OK |
| 1 | *S.* Brancaster | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Bredeney | O:4 | O3-4-21 | 20163 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Butantan | O:3,10 | O3-4-21 | 6195 | Probably *S.* Butantan | O:3 | TN | OK |
| 1 | *S.* Chester | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Coeln | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Colindale | O:7 | O7 | 15 | *Salmonella* | O:7 | TN | OK |
| 1 | *S.* Dublin | O:9 | O9 | 15 | *Salmonella* | O:9 | TN | OK |
| 1 | *S.* Durban | O:9 | O9 | 15 | *Salmonella* | O:9 | TN | OK |
| 1 | *S.* Eastbourne | O:9 | O9 | 15 | *Salmonella* | O:9 | TN | OK |
| 4 | *S.* Give | O:3,10 | O3-4-21 | 291165 | Probably *S.* Give | O:3 | TN | OK |
| 3 | *S.* Haifa | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Heidelberg | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 2 | *S.* Indiana | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 5 | *S.* Infantis | O:7 | O7 | 345 | *Salmonella* | O:7 | TN | OK |
| 1 | *S.* Ituri | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Javiana | O:9 | O9 | 15 | *Salmonella* | O:9 | TN | OK |
| 1 | *S.* Jerusalem | O:7 | O7 | 15 | *Salmonella* | O:7 | TN | OK |
| 1 | *S.* Kapemba | O:9 | O9 | 15 | *Salmonella* | O:9 | TN | OK |
| 1 | *S.* Lagos | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 2 | *S.* Liverpool | O:1,3,19 | O3-4-21 | 165 | *Salmonella* | O:3 | TN | OK |
| 2 | *S.* Llandoff | O:1,3,19 | O3-4-21 | 165 | *Salmonella* | O:3 | TN | OK |
| 6 | *S.* London | O:3,10 | O3-4-21 | 389865 | Probably *S.* London | O:3 | TN | OK |
| 4 | *S.* Meleagridis | O:3,10 | O3-4-21 | 4935 | *Salmonella* | O:3 | TN | OK |
| 1 | *S.* Mikawasima | O:7 | O7 | 15 | *Salmonella* | O:7 | TN | OK |
| 1 | *S.* Montevideo | O:7 | O7 | 15 | *Salmonella* | O:7 | TN | OK |
| 4 | *S.* Muenster | O:3,10 | O3-4-21 | 105 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Napoli | O:9 | O9 | 15 | *Salmonella* | O:9 | TN | OK |
| 1 | *S.* Nyborg | O:3,10 | O3-4-21 | 7455 | *Salmonella* | O:3 | TN | OK |
| 3 | *S.* Nyborg | O:3,10 | O3-4-21 | 105 | *Salmonella* | O:3 | TN | OK |
| 1 | *S.*Oranienburg | O:7 | O7 | 195 | *Salmonella* | O:7 | TN | OK |
| 1 | *S.* Panama | O:9 | O9 | 15 | *Salmonella* | O:9 | TN | OK |
| 11 | *S.* Paratyphi B dT+ | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Saintpaul | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Sandiego | O:4 | O3-4-21 | 18447 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Schwarzengrund | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Singapore | O:7 | O7 | 105 | *Salmonella* | O:7 | TN | OK |
| 1 | *S.* Stanley | O:4 | O3-4-21 | 30459 | *Salmonella* | O:4 | TN | OK |
| 3 | *S.* Stanleyville | O:4 | O3-4-21 | 30459 | *Salmonella* | O:4 | TN | OK |
| 1 | *S.* Tennessee | O:7 | O7 | 15 | *Salmonella* | O:7 | TN | OK |
| 1 | *S.* Thompson | O:7 | O7 | 15 | *Salmonella* | O:7 | TN | OK |
| 5 | *S.* Typhimurium | O:4 | O3-4-21 | 429 | *Salmonella* | O:4 | TN | OK |
| 2 | *S.* Uganda | O:3,10 | O3-4-21 | 7455 | *Salmonella* | O:3 | TN | OK |
| 5 | *S.* Virchow | O:7 | O7 | 15 | *Salmonella* | O:7 | TN | OK |
| 3 | *S.* Weltevreden | O:3,10 | O3-4-21 | 105 | *Salmonella* | O:3 | TN | OK |
| TOTAL:102 | | | | | Total TN: | | 102 | |
| | | | | | Total FP: | | 0 | |
| | | | | | Exclusivitiy (specificity): | | 100% | |

dT+: D-Tartrate fermenting isolates.
TN: True Negative, including probable results which are not complete identifications and need classical methods, i.e. slide-agglutination and biochemical test, to be confirmed; FP: False Positive.
[a] Obtained by classical methods, i.e. slide-agglutination and biochemical tests.
[b] The results presented here are those obtained using one of the 3 modules, depending on the serogroup of the analyzed isolate.

and 27 *S.* Rissen was correctly obtained by the DSS. The ability of the MOL-PCR O9 to detect the vaccine strains of *S.* Enteritidis was validated with the correct discrimination between 5 wild type *S.* Enteritidis coming from the field and 2 isolates coming each from the one of the commercial vaccines AviPro SALMONELLA VAC E and Salmovac SE. The serotype of 12 *S.* Gallinarum isolates was confirmed by the MOL-PCR O9 and they were correctly discriminated into 2 *S.* Gallinarum var. Gallinarum and 10 *S.* Gallinarum var. Pullorum. Unfortunately, not enough *S.* Choleraesuis and *S.* Gallinarum were available in the NRC collection to achieve the validation criteria of at least 25 *Salmonella* isolates per targeted seroytpes. As no false negative was obtained among the targeted *Salmonella* used here for the inclusivity tests, the sensitivity was determined to be 100% for each of the tested modules (Table 3 and Supplementary Table S4).

For the exclusivity tests, 134 untargeted *Salmonella* isolates were analyzed with the 3 modules, among which 102 belonged to one of the serogroups targeted by the method and were correctly clustered by the serogrouping probes (Table 4), and 32 belonged to 13 other serogroups

**Table 5**
Results of the exclusivity tests using non-*Salmonella* isolates and *Salmonella* isolates for which the serogroup cannot be determined by the developed method.

| Number of isolates tested | Reference Identification[a] | Luminex Identification result[b] | Comparison with the reference identification |
|---|---|---|---|
| 1 | *Acinetobacter baumanii* | No *Salmonella* | TN |
| 1 | *Bacillus cereus* | No *Salmonella* | TN |
| 1 | *Citrobacter koseri* | No *Salmonella* | TN |
| 1 | *Enterobacter aerogenes* | No *Salmonella* | TN |
| 1 | *Enterococcus faecium* | No *Salmonella* | TN |
| 2 | *Escherichia coli* | No *Salmonella* | TN |
| 1 | *Klebsiella oxytoca* | No *Salmonella* | TN |
| 1 | *Klebsiella pneumoniae* | No *Salmonella* | TN |
| 5 | *Listeria monocytogenes* | No *Salmonella* | TN |
| 1 | *Morganella morganii* | No *Salmonella* | TN |
| 4 | *Neisseria meningitidis* | No *Salmonella* | TN |
| 1 | *S.* Ago (O:30) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Agoueve (O:13,22) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Brive (O:42) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Carmel (O:17) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Cero (O:18) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Cotham (O:28) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Dugbe (O:45) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Durham (O:13,23) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Ebrie (O:35) | *Salmonella* - Unknown serogroup | TN |
| 3 | *S.* Gaminara (O:16) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Havana (O:13) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Hvittingfoss (O:16) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Idikan (O:13) | *Salmonella* - Unknown serogroup | TN |
| 2 | *S.* Kasenyi (O:38) | *Salmonella* - Unknown serogroup | TN |
| 2 | *S.* Kedougou (O:13) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Kisarawe (O:11) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Mgulani (O:38) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Monschaui (O:35) | *Salmonella* - Unknown serogroup | TN |
| 3 | *S.* Paratyphi A (O:2) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Pomona (O:28) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Poona (O:13,22) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Rubislaw (O:11) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Telelkebir (O:13,23) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Umbilo (O:28) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Urbana (O:30) | *Salmonella* - Unknown serogroup | TN |
| 1 | *S.* Wandsworth (O:39) | *Salmonella* - Unknown serogroup | TN |
| 1 | *Serratia odorifera* | No *Salmonella* | TN |
| 1 | *Shigella boydii* | No *Salmonella* | TN |
| 1 | *Shigella flexneri* | No *Salmonella* | TN |
| 1 | *Shigella sonei* | No *Salmonella* | TN |
| 1 | *Staphylococcus aureus* | No *Salmonella* | TN |
| 1 | *Staphylococcus epidermidis* | No *Salmonella* | TN |
| 1 | *Staphylococcus mileri* | No *Salmonella* | TN |
| 1 | *Staphylococcus saprophyticus* | No *Salmonella* | TN |
| 1 | *Streptococcus agalactiae* | No *Salmonella* | TN |
| 1 | *Streptococcus bovis* | No *Salmonella* | TN |
| 1 | *Streptococcus dysgalactiae* | No *Salmonella* | TN |
| 1 | *Streptococcus pyogenes* | No *Salmonella* | TN |
| 1 | *Vibrio alginolyticus* | No *Salmonella* | TN |
| 1 | *Yersinia enterocolitica* | No *Salmonella* | TN |
| Total:65 | | Total TN: | 65 |
| | | Total FP: | 0 |
| | | Exclusivitiy (specificity): | 100% |

[a] According to the NRC collection or obtained by classical method (i.e. slide-agglutination and biochemical tests) for the *Salmonella* isolates.
[b] Obtained with each of the 3 developed assays.

and were correctly reported as "Unknown serogroup" by the DSS (Table 5). Additionally, for each module, 1 *Salmonella* isolate per serotype targeted by the 2 other assays was analyzed and correctly determined as "Unknown serogroup" (data not shown). For these cases described above, the DSS recommended to use the classical method to obtain the complete identification. Finally, 33 non-*Salmonella* isolates were analyzed with the 3 MOL-PCR assays and successfully reported as "No *Salmonella*" by the DSS (Table 5). From these exclusivity tests, no false positive was obtained and the specificity was determined to be 100%. It could be observed that the 3 MOL-PCR assays were even able to make "probable" serotype predictions, thereby recommending the user for these cases what classical test to perform for completing the identification. These predictions concerned the serotypes: Butantan, Give and London (Table 4).

In conclusion, based on the data produced during this comparison study with the inclusivity (Table 3) and exclusivity tests (Tables 4 and 5), the accuracy of the 3 new MOL-PCR assays was determined to be 100%.

### 3.4. Modification of the MOL-PCR BASE

During the development of the 3 new MOL-PCR modules, in view of combining the 4 modules into one complete genoserotyping system (Fig. 2), the MOL-PCR BASE module (Gand et al., 2020) was modified at two levels.

First, an adaptation was needed because when this assay was implemented for routine analyses at the Belgian NRC in 2019 and performed in parallel with the classical method, 3 *S.* Haifa were incorrectly

**Table 6**
Probes involved in the identification of S.Paratyphi B (dT-/dT+) using the adapted MOL-PCR BASE module.

| Reference identification[a] | Number of tested isolates | Prime numbers: | 3 | 5 | 7 | 83 | 13 | 11 | 17 | GPP | Comparison expected vs. obtained results |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Antigenic formula[b] | invA (*Salmonella*) | SAL-73 (H:1,2) | STID16 (O:4) | STID333 (Paratyphi B) | STID334 (Paratyphi B) | STID71 (Paratyphi B)[c] | STID34 (dT- variant) | | |
| *S.* Paratyphi B dT- | 13 | 1,4,[5],12:b:1,2 | + | + | + | + | + | + | + | 2118165 | 13 TP |
| *S.* Paratyphi B dT+ | 16 | 1,4,[5],12:b:1,2 | + | + | + | + | + | + | – | 1246245 | 16 TP |
| *S.* Paratyphi B dT+ | 1 | 1,4,[5],12:b:1,2 | + | + | + | + | + | – | – | 113295 | 1 TP |
| *S.* Paratyphi B dT+ | 37 | 1,4,[5],12:b:1,2 | + | + | + | – | + | + | – | 15015 | 37 TP |
| *S.* Haifa | 3 | 1,4,[5],12:$z_{10}$:1,2 | + | + | + | – | + | – | – | 1365 | 3 TN |
| Total | 70 | | | | | | | | Total TP: | | 67 |
| | | | | | | | | | Total TN: | | 3 |
| | | | | | | | | | Total FP: | | 0 |
| | | | | | | | | | Total FN: | | 0 |
| | | | | | | | | | Inclusivity (sensitivity): | | 100% |
| | | | | | | | | | Exclusivity (specificity): | | 100% |
| | | | | | | | | | Accuracy: | | 100% |

TP: True Positive; TN: True Negative; FP: False Positive; FN: False Negative.
[a] Obtained with the classical method, i.e. slide-agglutination and biochemical test.
[b] According to Grimont and Weill (2007)..
[c] Present in 54% of the S. Paratyphi B population in EnteroBase (Gand et al., 2020).

confounded with an uncommon genotype of *S.* Paratyphi B var. Java dT+, detected with the GPP1365. This GPP is obtained with the rare probe combination including invA (*Salmonella*), SAL-73 (H:1,2), STID16 (O:4) and STID334 (Paratyphi B) but without STID71 (*hemD*22) (Table 6). Indeed, this probe combination was retrieved in only 1 of the 54 *S.* Paratyphi B dT + tested, despite the fact that the SNP marker of *hemD*22 is absent in 46% of the *S.* Paratyphi B population in EnteroBase. No similar problem was observed with the *S.* Paratyphi B dT- isolates (16) as they are always positive for the probe STID34, which discriminate them from the *S.* Haifa isolates that are dT+. To avoid any confusion between *S.* Haifa and *S.* Paratyphi B dT+, the probe STID333, targeting the marker *SPAB_01124* described by Zhai et al. (2014) for the detection of *S.* Paratyphi B (Table 2), was added to the MOL-PCR BASE with the prime number 83. It appeared that the marker *SPAB_01124* was absent in some *S.* Paratyphi B dT + isolates, but not at the same time than *hemD*22. Consequently, the specific detection was successfully obtained for the 54 *S.* Paratyphi B dT + tested, with the combination of STID334 with at least STID333 or STID71 (Table 6). Concerning the 3 *S.* Haifa, the marker *SPAB_01124* was not present in these isolates and they were reported as *Salmonella* O:4 by the DSS (Supplementary Table S4).

Secondly, the serogrouping probe STID31 (targeting the serogroup O:3,10) was replaced by the probe STID301 to detect all the *Salmonella* isolates belonging to the serogroup O:3, including O:3,10 and O:1,3,19. All the tested *Salmonella* isolates belonging to O:3 (56 to O:3,10 and 32 to O:1,3,19) were correctly serogrouped as O:3 by STID301 (Supplementary Table S4).

To test the results' interpretation and the recommendations provided by the DSS, all the isolates (553) used in this study were analyzed with the module BASE. From this, 94 were completely identified with the module BASE only and 394 were partially identified (i.e. probable serotype or determination of the serogroup only), 32 were determined to be *Salmonella* from unknown serogroup and 33 were identified as No *Salmonella*, all in agreement with the expected results. For the 394 *Salmonella* isolates partially identified, 389 were successfully recommended by the DSS to be analyzed by one of the 3 other MOL-PCR assay. To confirm some of the probable serotype, the DSS recommended to complete the identification using the classical method (Supplementary Table S4). The MOL-PCR BASE was already validated by Gand et al. (2020) with an accuracy of 99.7%, so its specificity was only evaluated here for the addition of STID333 and STID301. As no false positive nor false negative were obtained (Table 6 and Supplementary Table S4), the accuracy was determined to be 100% for these modifications.

## 4. Discussion

In the present study, 3 new MOL-PCR assays (MOL-PCR O3-4-21, MOL-PCR O7 and MOL-PCR O9) were developed for a fast and accurate genoserotyping of common *Salmonella* serotypes (and their variants) which are possible to be isolated in Belgium from the poultry and pork sector. The MOL-PCR assays O7 and O9 can also be used for a fast detection of important invasive serotypes (e.g. *S.* Choleraesuis and *S.* Gallinarum) if, based on clinical symptoms, they are suspected to cause infections in animal breeding. As such, actions can be quickly taken to avoid the spread of these invasive serotypes. Moreover, when the sampling at the poultry farms is too close to the time point of vaccination of the breeding against *S.* Enteritidis, the MOL-PCR O9 can be used for a reliable discrimination between wild type and vaccine strains. For the validation of these 3 tests, a comparison study with the classical method, using 464 bacterial isolates, was conducted and an accuracy of 100% was obtained for the detection of the serotypes included in the modules, i.e. Agona, Anatum, Brandenburg, Choleraesuis, Derby, Enteritidis (including the vaccine strains), Gallinarum var. Gallinarum/Pullorum, Livingstone, Mbandaka, Minnesota, Ohio, Rissen and Senftenberg.

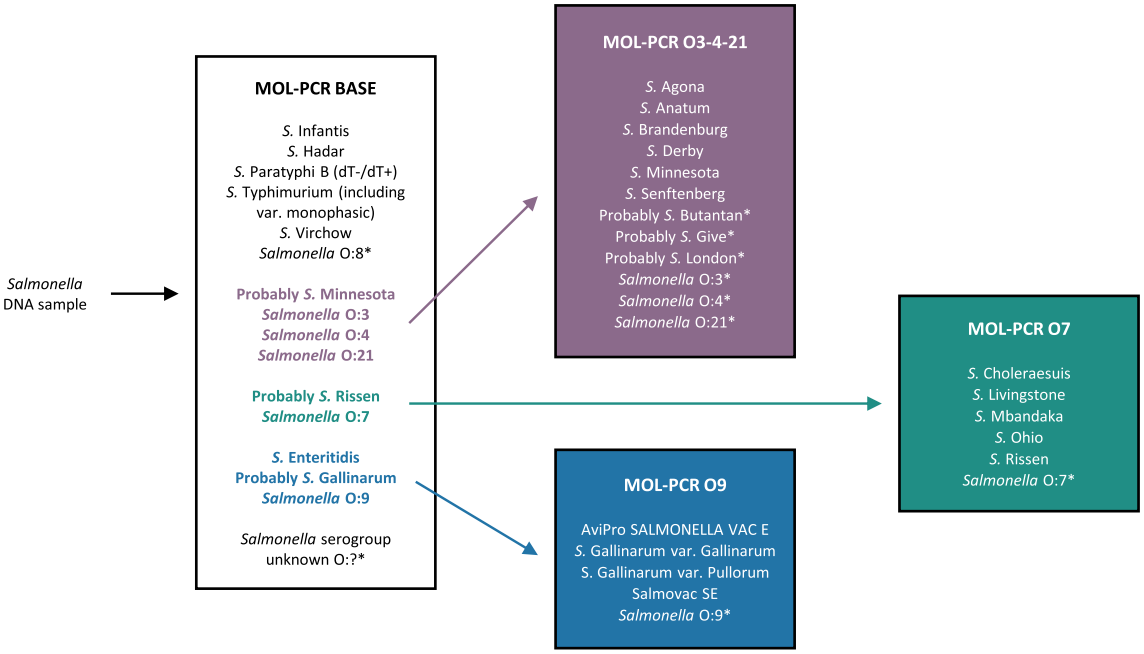


**Fig. 2.** Recommended workflow for *Salmonella* serotype identification using the 4 MOL-PCR assays, as configured in the DSS. The figure shows the serogroups, serotypes and their variant that can be identified with each of the 4 modules. The MOL-PCR BASE is used for a first screening of all DNA samples extracted from *Salmonella* isolates. For the partial identification obtained using this module, indicated in purple, green and blue in the figure, the DSS recommends to use the MOL-PCR O3-4-21, MOL-PCR O9 or MOL-PCR O7, respectively, to complete the identification if possible. For the partial identifications marked with a star (*), the DSS indicates this time to use the classical method, i.e. slide-agglutination and biochemical tests, to complete the identification. . (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Additionally, a previously developed MOL-PCR assay (Gand et al., 2020), called here module BASE, was modified. This module BASE can identify the 6 serotypes (Enteritidis, Hadar, Infantis, Paratyphi B including its variant Java, Typhimurium including its monophasic variant and Virchow) subjected to an official control (EU regulation N°2160/2003, Belgian royal decree 27/04/2007 and Belgian FASFC note BP-FDS/LABO/1470050 v7) and cluster the other unknown serotypes in one of the common serogroup (O:3, O:4, O:7, O:8, O:9, O:21) if they belong to one of them. Despite the fact that the module BASE was previously validated by (Gand et al., 2020) with the analysis of 1004 bacterial isolates belonging to 114 of the most common serotypes, 3 rare *S.* Haifa were wrongly reported in the current study as a in Belgium uncommon genotype of Paratyphi B var. Java, thus leading to a false positive result. This shows the limitation of the tests performed for the validation of alternative molecular methods which use only isolates belonging to the most common serotypes (*S.* Haifa was isolated only 7 times in Belgium during the last 5 years, personal communication NRC) because it would be too labor intensive to analyze the 2500 *Salmonella* serotypes of the KWL scheme. Moreover, only few laboratories in the world have the complete collection of these 2500 serotypes. In the current study, modifications were made to the BASE module to exclude the 3 *S.* Haifa isolates detected as false positive. Additionally, improvements were made to the detection of *Salmonella* isolates belonging to serogroup O:3, including the cluster O:1,3,19. These 2 modifications were validated, identically as for the 3 new MOL-PCR assays, with an accuracy of 100%. Finally, a DSS accessible through a web-application was created for an automatic interpretation of the Luminex data, using a barcode system (GPP), and for the centralization of the results in a database improving the surveillance at a national level. Furthermore, this DSS is also able to provide recommendations to the users in case of partial identifications, i.e. probable serotype or serogroup clustering only.

For the serotypes Derby, Paratyphi B (without *hemD22* marker), and Senftenberg, the genotype distribution observed during the study, i.e. 34%, less than 1% and 25%, respectively, was different from the one estimated *in silico* with EnteroBase, i.e. 16%, 46% and 10%, respectively. This can be explained by the fact that the *Salmonella* isolates used in this study were selected from the collection of the Belgian NRC, composed of routine samples isolated in Belgium, and therefore are more representative of the genotypes circulating in this country. In contrast, the genomic data of *Salmonella* samples uploaded in EnteroBase come from all over the world but are not identical to the frequencies of prevalent serotypes obtained with a national surveillance program.

Altogether, the 4 MOL-PCR assays piloted by the DSS compose a validated *Salmonella* genoserotyping system. The MOL-PCR BASE module is recommended to be used for a first screening of new samples because if one of the serotypes targeted by this module is present at the breeding site, strict and constraining disinfection procedures must be performed at the farm, and the animals are excluded from the food chain which leads to economical loss for the farmers. In case of a partial identification result obtained with this first assay, the DSS displays a recommendation to the user on which of the 3 other MOL-PCR modules (MOL-PCR O3-4-21, MOL-PCR O7 and MOL-PCR O9) he/she can perform to complete the identification of the isolates (Fig. 2). In some cases, the web-application will recommend to directly switch to the serotyping by slide-agglutination with targeted antisera to test.

The MOL-PCR and the Luminex technology used to develop the genoserotyping system presented in this study allow a high-throughput analysis as the method is based on experiments in a 96-well plate. Additionally, the Mag-Pix apparatus offers the possibility to perform several MOL-PCR assays (including other tests based on the Luminex xTAG technology) at the same time, thereby saving time and money. Indeed, the MOL-PCR BASE module can be run for new samples simultaneously with the other MOL-PCR modules used to complete the identification of isolates partially identified the day before, with a limit of 96 samples in total. The addition of samples to the plate does not drastically increase the price of the analysis, but rather allows to reduce the cost per sample (Gand et al., 2020). Compared to the classical method (i.e. slide-agglutination and biochemical tests) for which 2–9

days of analyses are usually needed, performed by an experimented technician at the NRC, the developed genoserotyping system does not require particular skills, and objective results are obtained in 1–2 days for the targeted serotypes. Therefore, the genoserotyping test can easily be implemented in first line laboratories as well as in NRCs and helps to reduce the analysis time, thus complying with the short TAT required by the food sector. Moreover, as the price of the antisera is constantly increasing, the use of this alternative molecular method is also cost-effective. The professionals of the food sector are not ready to pay the expensive price required to completely identify by classical method the *Salmonella* serotypes which are not subjected to an official control. However, when the new and less expensive MOL-PCR assays developed in this study will be used, they will be more disposed to do so. Additionally, as all the identification results obtained with these assays are centralized in a database included in the DSS, this will help the transmission of the serotype identification data to the NRC and thus improve the *Salmonella* surveillance at a national level.

According to the serotyping analyses performed at the NRC between 2017 and 2018, the MOL-PCR BASE could have completely identified 59% and 54% of the *Salmonella* isolates coming from food and veterinary (including animal feed) sources, respectively. Among the remaining unidentified samples, 36% from food source and 50% from veterinary source (including animal feed) could have been genoserotyped by one of the 3 other modules. Concerning the rest of the samples, the serotype identification must be determined using the classical method. But the MOL-PCR and Luminex technologies are modular and the composition of the modules can easily be adapted if needed, like it was the case for the module BASE in this study, following the evolution of the most common serotypes circulating in Belgium according to the database of the DSS or the modifications of the law. Since 2017, *Salmonella* serovar Newport reached the top 5 of the most commonly reported cases in Europe (EFSA, 2019, 2018). A part of the genotypes composing the *S.* Newport population can already be detected as “Probable serotype” by the module BASE (Gand et al., 2020). But the detection of this serotype could be improved by including complementary markers either in the module BASE or in the module O7 (thus becoming the module O7-8).

Another alternative method for *Salmonella* typing is based on WGS which offers a complete identification of *Salmonella* isolates (including serotype, variant and subtype), in addition to providing other information such as antibiotic resistance or phylogenetic profiles (Ibrahim and Morin, 2018; Pornsukarom et al., 2018; Yachison et al., 2017). WGS is already routinely used by big public health institutes in Canada, France, United Kingdom and U.S.A. (Allard, 2016; Ashton et al., 2016; Institut Pasteur, 2018; Jain et al., 2019). But despite the fact that this technology is more complete compared to target-based methods like MOL-PCR, it is time-consuming (1 analysis takes 3–5 days) and too expensive for small institutions with limited resources, such as first line laboratories which have to respect short TAT and cannot wait for sample batching to reduce analysis costs (Ibrahim and Morin, 2018).

In conclusion, unless WGS will become the mandatory, less expensive and more rapid gold standard method in the future for the characterization of *Salmonella* (including the serotyping), target-based molecular methods such as MOL-PCR linked to the Luminex technology still have their utility. The genoserotyping system developed in this study is able to perform a fast and cheap identification of the most common *Salmonella* serotypes isolated from the poultry and pork sectors. With this method, objective and accurate results are obtained thanks to the automatic interpretation of the Luminex data by a DSS which can also give recommendations for further testing in case of partial identification. Consequently, the method is fully adapted to the needs of the food-producing animal sector. The database present in the DSS will also help to improve the surveillance of *Salmonella* serotypes at the national level and orientate the future modifications of the module composition to follow the trends of the most prevalent serotypes in Belgium.

## Declaration of competing interest

All authors declare that they have no conflict of interest.

## Acknowledgements

The research that yielded these results was funded by the Belgian Federal Public Service of Health, Food Chain Safety and Environment through the contract RT 14/5 SALMSTID 1. Whole genome sequencing was performed at the service Transversal Activities in Applied Genomics and the DSS was developed in collaboration with the ICT service, both at Sciensano. The authors want to thank the technicians of the National Reference Center for *Salmonella* and *Shigella*, for their expertise in serotyping by slide-agglutination and their participation to the comparison study; Sylvia Broeders from the service Quality of the Laboratories for providing the non-*Salmonella* strains; and the technicians of the service Transversal Activities in Applied Genomics for their technical support of the MagPix.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fm.2020.103534.

## References

Achtman, M., Wain, J., Weill, F.X., Nair, S., Zhou, Z., Sangal, V., Krauland, M.G., Hale, J.L., Harbottle, H., Uesbeck, A., Dougan, G., Harrison, L.H., Brisse, S., 2012. Multilocus sequence typing as a replacement for serotyping in Salmonella enterica. PLoS Pathog. 8. https://doi.org/10.1371/journal.ppat.1002776.

Ågren, J., Sundström, A., Håfström, T., Segerman, B., 2012. Gegenees: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. PloS One 7. https://doi.org/10.1371/journal.pone.0039107.

Alikhan, N.F., Zhou, Z., Sergeant, M.J., Achtman, M., 2018. A genomic overview of the population structure of Salmonella. PLoS Genet. 14, 1–13. https://doi.org/10.1371/journal.pgen.1007261.

Allard, M.W., 2016. The future of whole-genome sequencing for public health and the clinic. J. Clin. Microbiol. 54, 1946–1948. https://doi.org/10.1128/JCM.01082-16.

Alves Batista, D.F., de Freitas Neto, O.C., de Almeida, A.M., Maboni, G., de Carvalho, T.F., de Carvalho, T.P., Barrow, P.A., Berchieri, A., 2018. Evaluation of pathogenicity of Salmonella Gallinarum strains harbouring deletions in genes whose orthologues are conserved pseudogenes in S. Pullorum. PLoS One 13, 1–18. https://doi.org/10.1371/journal.pone.0200585.

Antunes, P., Mourão, J., Campos, J., Peixe, L., 2016. Salmonellosis: the role of poultry meat. Clin. Microbiol. Infect. 22, 110–121. https://doi.org/10.1016/j.cmi.2015.12.004.

Ashton, P.M., Nair, S., Peters, T.M., Bale, J.A., Powell, D.G., Painset, A., Tewolde, R., Schaefer, U., Jenkins, C., Dallman, T.J., de Pinna, E.M., Grant, K.A., 2016. Identification of *Salmonella* for public health surveillance using whole genome sequencing. PeerJ 4, e1752. https://doi.org/10.7717/peerj.1752.

Batista, D.F.A., de Freitas Neto, O.C., Lopes, P.D., de Almeida, A.M., Barrow, P.A., Berchieri, A., 2013. Polymerase chain reaction assay based on ratA gene allows differentiation between Salmonella enterica subsp. enterica serovar Gallinarum biovars Gallinarum and Pullorum. J. Vet. Diagn. Invest. 25, 259–262. https://doi.org/10.1177/1040638713479361.

Demirbilek, S.K., 2016. Salmonellosis in animals. In: Mascellino, M.T. (Ed.), Salmonella - A Re-emerging Pathogen. IntechOpen, pp. 18. https://doi.org/10.5772/57353.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. https://doi.org/10.1093/nar/gkh340.

EFSA, 2019. The European union one health 2018 zoonoses report. EFSA Journal. https://doi.org/10.2903/j.efsa.2019.5926.

EFSA, 2018. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. EFSA Journal. https://doi.org/10.2903/j.efsa.2018.5500.

Franklin, K., Lingohr, E.J., Yoshida, C., Anjum, M., Bodrossy, L., Clark, C.G., Kropinski, A.M., Karmali, M.A., 2011. Rapid genoserotyping tool for classification of Salmonella serovars. J. Clin. Microbiol. 49, 2954–2965. https://doi.org/10.1128/JCM.02347-10.

Gand, M., Mattheus, W., Roosens, N.H.C., Dierick, K., Marchal, K., De Keersmaecker, S.C.J., Bertrand, S., 2020. A multiplex oligonucleotide ligation-PCR method for the genoserotyping of common Salmonella using a liquid bead suspension assay. Food Microbiol. 87. https://doi.org/10.1016/j.fm.2019.103394.

Gand, M., Mattheus, W., Saltykova, A., Roosens, N., Dierick, K., Marchal, K., De Keersmaecker, S.C.J., Bertrand, S., 2019. Development of a real-time PCR method for the genoserotyping of Salmonella Paratyphi B variant Java. Appl. Microbiol. Biotechnol. https://doi.org/10.1007/s00253-019-09854-4.

Griffin, P.M., O'Brien, S.J., 2013. The “decline and Fall” of nontyphoidal salmonella in the United Kingdom. Clin. Infect. Dis. 56, 705–710. https://doi.org/10.1093/cid/cis967.
Grimont, P.A.D., Weill, F.-X., 2007. Antigenic Formulae of the Salmonella Serovars, ninth ed. Institut Pasteur, Paris.
Hayward, M.R., Jansen, V.A.A., Woodward, M.J., 2013. Comparative genomics of Salmonella enterica serovars Derby and Mbandaka, two prevalent serovars associated with different livestock species in the UK. BMC Genom. 14, 1. https://doi.org/10.1186/1471-2164-14-365.
Heredia, N., García, S., 2018. Animals as sources of food-borne pathogens: a review. Anim. Nutr. 4, 250–255. https://doi.org/10.1016/J.ANINU.2018.04.006.
Herrera-León, S., McQuiston, J.R., Usera, M.A., Fields, P.I., Garaizar, J., Echeita, M.A., 2004. Multiplex PCR for distinguishing the most common phase-1 flagellar antigens of Salmonella spp. J. Clin. Microbiol. 42, 2581–2586. https://doi.org/10.1128/JCM.42.6.2581-2586.2004.
Ibrahim, G.M., Morin, P.M., 2018. Salmonella serotyping using whole genome sequencing. Front. Microbiol. 9, 2993. https://doi.org/10.3389/fmicb.2018.02993.
Institut Pasteur, 2018. Rapport D'activité Annuel 2018 Année D'exercice 2017.
Jain, S., Mukhopadhyay, K., Thomassin, P.J., 2019. An economic analysis of salmonella detection in fresh produce, poultry, and eggs using whole genome sequencing technology in Canada. Food Res. Int. 116, 802–809. https://doi.org/10.1016/j.foodres.2018.09.014.
Jean-Gilles Beaubrun, J., Ewing, L., Jarvis, K., Dudley, K., Grim, C., Gopinath, G., Flamer, M.L., Auguste, W., Jayaram, A., Elmore, J., Lamont, M., McGrath, T., Hanes, D.E., 2014. Comparison of a PCR serotyping assay, Check&Trace assay for Salmonella, and Luminex Salmonella serotyping assay for the characterization of Salmonella enterica identified from fresh and naturally contaminated cilantro. Food Microbiol. 42, 181–187. https://doi.org/10.1016/j.fm.2014.02.008.
Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054.
Liang, D.W., Lu, J.H., Wu, Q., Ke, B.X., Jiang, C.H., Long, J., Fang, Y.P., Lin, L.J., Zeng, N.Y., Fu, L., Jiang, L.X., 2016. Comparing the ability of luminex xMAP® salmonella serotyping assay and traditional serotyping method for serotyping salmonella isolated from southern Chinese population. J. Appl. Microbiol. 120, 1668–1676. https://doi.org/10.1111/jam.13106.
Majowicz, S.E., Musto, J., Scallan, E., Angulo, F.J., Kirk, M., O'Brien, S.J., Jones, T.F., Fazil, A., Hoekstra, R.M., 2010. The global burden of nontyphoidal *Salmonella* gastroenteritis. Clin. Infect. Dis. 50, 882–889. https://doi.org/10.1086/650733.
Maurischat, S., Szabo, I., Baumann, B., Malorny, B., 2015. Rapid real-time PCR methods to distinguish Salmonella Enteritidis wildtype field isolates from vaccine strains Salmovac SE/Gallivac SE and AviPro SALMONELLA VAC. E. J. Microbiol. Methods 112, 92–98. https://doi.org/10.1016/j.mimet.2015.03.015.
Oxford Analytica, 2012. A Report Produced for the International Federation for Animal Health: the Costs of Animal Disease. https://doi.org/10.1057/9780230358522_6.
Pedersen, K., Sørensen, G., Löfström, C., Leekitcharoenphon, P., Nielsen, B., Wingstrand, A., Aarestrup, F.M., Hendriksen, R.S., Baggesen, D.L., 2015. Reappearance of Salmonella serovar Choleraesuis var. Kunzendorf in Danish pig herds. Vet. Microbiol. 176, 282–291. https://doi.org/10.1016/j.vetmic.2015.01.004.
Pornsukarom, S., van Vliet, A.H.M., Thakur, S., 2018. Whole genome sequencing analysis of multiple Salmonella serovars provides insights into phylogenetic relatedness, antimicrobial resistance, and virulence markers across humans, food animals and agriculture environmental sources. BMC Genom. 19, 801. https://doi.org/10.1186/s12864-018-5137-4.
Ryan, M.P., O'Dwyer, J., Adley, C.C., 2017. Evaluation of the complex nomenclature of the clinically and veterinary significant pathogen Salmonella. BioMed Res. Int. https://doi.org/10.1155/2017/3782182. 2017.
Van Den Bulcke, M., Lievens, A., Barbau-Piednoir, E., Mbongolombella, G., Roosens, N., Sneyers, M., Casi, A.L., 2010. A theoretical introduction to “combinatory SYBR®Green qPCR Screening”, a matrix-based approach for the detection of materials derived from genetically modified plants. Anal. Bioanal. Chem. 396, 2113–2123. https://doi.org/10.1007/s00216-009-3286-7.
Van den Bulcke, M., Lievens, A., Leunda, A., MbongoloMbella, E., Barbau-Piednoir, E., Sneyers, M., 2008. Transgenic Plant Event Detection. Patent WO 2008/092866.
Wuyts, V., Mattheus, W., Roosens, N.H.C., Marchal, K., Bertrand, S., De Keersmaecker, S.C.J., 2015. A multiplex oligonucleotide ligation-PCR as a complementary tool for subtyping of Salmonella Typhimurium. Appl. Microbiol. Biotechnol. 99, 8137–8149. https://doi.org/10.1007/s00253-015-6831-7.
Yachison, C.A., Yoshida, C., Robertson, J., Nash, J.H.E., Kruczkiewicz, P., Taboada, E.N., Walker, M., Reimer, A., Christianson, S., Nichani, A., Nadon, C., 2017. The validation and implications of using whole genome sequencing as a replacement for traditional serotyping for a national Salmonella reference laboratory. Front. Microbiol. 8, 1–9. https://doi.org/10.3389/fmicb.2017.01044.
Yoshida, C., Gurnik, Simone, Ahmad, A., Blimkie, T., Murphy, S.A., Kropinski, A.M., Nash, J.H.E., 2016a. Evaluation of molecular methods for identification of. J. Clin. Microbiol. 54https://doi.org/10.1128/JCM.00262-16. 1992–1998.
Yoshida, C.E., Kruczkiewicz, P., Laing, C.R., Lingohr, E.J., Gannon, V.P.J., Nash, J.H.E., Taboada, E.N., 2016b. The salmonella in silico typing resource (SISTR): an open web-accessible tool for rapidly typing and subtyping draft salmonella genome assemblies. PloS One 11, 1–17. https://doi.org/10.1371/journal.pone.0147101.
Zhai, L., Yu, Q., Bie, X., Lu, Z., Lv, F., Zhang, C., Kong, X., Zhao, H., 2014. Development of a PCR test system for specific detection of Salmonella Paratyphi B in foods. FEMS Microbiol. Lett. 355, 83–89. https://doi.org/10.1111/1574-6968.12443.
Zhang, S., Yin, Y., Jones, M.B., Zhang, Z., Kaiser, B.L.D., Dinsmore, B.A., Fitzgerald, C., Fields, P.I., Deng, X., 2015. Salmonella serotype determination utilizing high-throughput genome sequencing data. J. Clin. Microbiol. 53, 1685–1692. https://doi.org/10.1128/JCM.00323-15.

## Glossary

*AC:* Allele call
*CSV:* Comma Separated Value
*CTRL_+ and CTRL_:* positive and negative controls
*DSS:* Decision Support System
*eBG:* eBurst group
*FASFC:* Federal Agency for the Security of the Food Chain
*GPP:* Gödel Prime Product
*KWL:* Kauffman-White-Le Minor
*MFI:* Medium Fluorescence Intensity
*MLST:* MultiLocus Sequence Typing
*MOL-PCR:* Multiplex Oligonucleotide Ligation – Polymerase Chain Reaction
*NRC:* National Reference Center
*SAPE:* Streptavidin-R-Phycoerythrin
*SISTR: Salmonella In Silico* Typing Resource
*SNR:* Signal to Noise Ratio
*SRA:* Sequence Read Archive
*ST:* Sequence Type
*TAT:* Turn-Around Time
*WGS:* Whole Genome Sequencing