General Enquiries on the form should be made to:
Defra, Procurements and Commercial Function (Evidence Procurement Team)
E-mail: research.competitions@defra.gsi.gov.uk

Evidence Project Final Report

- **Note**
  In line with the Freedom of Information Act 2000, Defra aims to place the results of its completed research projects in the public domain wherever possible.
  The Evidence Project Final Report is designed to capture the information on the results and outputs of Defra-funded research in a format that is easily publishable through the Defra website
  An Evidence Project Final Report must be completed for all projects.
- This form is in Word format and the boxes may be expanded, as appropriate.
- **ACCESS TO INFORMATION**
  The information collected on this form will be stored electronically and may be sent to any part of Defra, or to individual researchers or organisations outside Defra for the purposes of reviewing the project. Defra may also disclose the information to any outside organisation acting as an agent authorised by Defra to process final research reports on its behalf. Defra intends to publish this form on its website, unless there are strong reasons not to, which fully comply with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.
  Defra may be required to release information, including personal data and commercial information, on request under the Environmental Information Regulations or the Freedom of Information Act 2000. However, Defra will not permit any unwarranted breach of confidentiality or act in contravention of its obligations under the Data Protection Act 1998. Defra or its appointed agents may use the name, address or other details on your form to contact you in connection with occasional customer research aimed at improving the processes through which Defra works with its contractors.

Project identification

1. Defra Project code: OZ0333
2. Project title: Comparative Evaluation of Molecular-Based Microarray Tools
3. Contractor organisation(s): AHVLA
   Woodham lane, Newhaw
   Surrey
4. Total Defra project costs (agreed fixed price): £ 317,692
5. Project: start date 01/11/09
   end date 31/03/12

6. It is Defra's intention to publish this form.

   Please confirm your agreement to do so..... **YES** ☐

   (a) When preparing Evidence Project Final Reports contractors should bear in mind that Defra intends that they be made public. They should be written in a clear and concise manner and represent a full account of the research project which someone not closely associated with the project can follow.

   Defra recognises that in a small minority of cases there may be information, such as intellectual property or commercially confidential data, used in or generated by the research project, which should not be disclosed. In these cases, such information should be detailed in a separate annex (not to be published) so that the Evidence Project Final Report can be placed in the public domain. Where it is impossible to complete the Final Report without including references to any sensitive or confidential data, the information should be included and section (b) completed. NB: only in exceptional circumstances will Defra expect contractors to give a "No" answer.

   In all cases, reasons for withholding information must be fully in line with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.

   (b) If you have answered NO, please explain why the Final report should not be released into public domain

Executive Summary

7. The executive summary must not exceed 2 sides in total of A4 and should be understandable to the intelligent non-scientist. It should cover the main objectives, methods and findings of the research, together with any other significant events and options for new work.

Salmonella spp. are zoonotic pathogens of significant public and animal health concern. The genus is comprised of the two species *S. bongori* and S. enterica. The later is further divided into six subspecies: S.*enterica*, S.*salamae*, S.*arizonae*, S.*diarizonae*, S.*indica* and S.*houtenae*. These can be further divided into serotypes based on the based on the presence of surface O (somatic) and H (flagellar) antigens. The White-Kauffmann-Le Minor scheme (Grimont and Weill, 2007) describes the antigenic formulae for all recognised *Salmonella* serotypes and is the international gold standard method for the serovar classification of *Salmonella* isolates. The classical method for determining the antigenic formula of a *Salmonella* isolate employs polyclonal antisera to determine the O (somatic) and H (flagellar) antigenic epitopes. The majority of *Salmonella* infections in humans and animals are caused by serotypes of *Salmonella enterica* subspecies *enterica*. There are differences in the pathogenicity, fitness, virulence, and host range of different *Salmonella* serotypes. The determination of *Salmonella* serotype is important for epidemiological and legislative purposes. The Kauffmann-White method for confirming and serotyping *Salmonella* is based on the detection of surface antigens. Using a range of antisera raised in animals, a series of slide agglutinations are performed. Several hurdles have limited the wider use of this serotyping method in tracing and eliminating *Salmonella*. These include: time-consuming testing, taking at least several days to obtain an end result; extensive training and expertise required to perform testing; up to 10-15% of cases in the EU (1% in the UK) are partly- or non-typable *Salmonella*, yielding inconclusive results; use of antisera which have limited shelf-life and that may be of varying quality.

The objective of this project was to evaluate alternative molecular methods of *Salmonella* detection with the aim of providing evidence for a rapid *Salmonella* alerting system. At the beginning of the project there were limited commercially available choices and the Premitest (PT) array was chosen for initial assessment. The PT array showed promising results using an easy to operate system that required no previous knowledge of *Salmonella*. However due to a non-compliance within the assessment of the PT array to detect to the accuracy of the current EU regulations, which state that the accuracy of detection of S. Enteritidis and S. Typhimurium must be >99.9% (being the most pathogenic serotypes) and that all other serotypes must be >99%, additional arrays were sought for evaluation for full development in to an effective rapid alerting system for *Salmonella*. The full report

for the PT array is in the annual report 2010.

Three arrays were sourced: a luminex based array (LUM array); a linear probe based array (SGSA) and an improved SNP based array (Check & Trace). The LUM & SGSA array were both in developmental stages and the Check & Trace array was an improved version of the Premitest array that was assessed originally . The C&T array originally marketed through DSM as Premitest was back in the control of the original developers of the array and they had made many improvements including automation and accreditation. A DNA library has been created which contains 104 *Salmonella* of known serotypes, additional non-*Salmonella* DNA and 2000+ routine submission samples. The 104 known *Salmonella* serotypes is a panel collated by the AHVLA over a number of years and has been used extensively in the development of different technologies at many collaborating institutes, including the PT array, the C&T array, the Luminex array and the SGSA in their initial development. As such each isolate is definitive in serotype and their intra and inter array capabilities can be directly compared. The panel includes the top 40 serotypes in the UK plus some challenging unusual rare serotypes and non-*Salmonella* isolates. The same colony from each isolate used will be assessed by the arrays to enable an exact comparison and evaluation.

The LUM array was being developed by the CDC laboratories in Atlanta and had undergone a preliminary evaluation in an European laboratory, which had reported excellent preliminary results. The array was already in an automated format and had the advantage of database updating to detect new and emerging serotypes rather than array changes and validation. Initial assessments using kits and quality control DNA provided by the CDC showed promise and further panels of AHVLA DNA were also assessed. The LUM array did not perform well in intra and inter assay validation on prepared DNA panels and one antigen consistently underperformed. There were also initial teething problems in setting up the system, some of which were resolved by the updating of the Luminex machine software and a better resourcing of reagents. Some of the problems proved insurmountable even with the full help and support of the CDC and it was determined that the array was prone to operator error with ease and lacked the robustness required by a rapid alerting system. In addition the stage of array development meant that all results were analysed and interpreted by the operator rather than software and as the array was still under development software was unlikely to be developed in the near future. This also limits the interpretation of results to personnel highly familiar with *Salmonella* structure. The technology has shown itself to be very demanding, requiring considerable effort to become fully proficient in its use.

The *Salmonella* Genotyping Array (SGSA) previously referred to in the original project proposal as the Identibac array has been collaboratively developed with the AHVLA, the Public Health Agency of Canada (PHAC) and the Austrian Institute of Technology. The SGSA took a long time to reach the validation stage due to problems with expanding the panel of probes within the array causing a knock on effect with the existing probes. This has now been resolved but remains a potential future problem when expanding the array to include additional or emerging serotypes. Current regulatory requirements require robust sensitivity and specificity for *Salmonella* serotyping to ensure high confidence in the correct identification serovars of major public and animal health performance. The SGSA generates an antigenic formula consistent with the White-Kauffmann-Le Minor scheme and currently includes the determination of 59 of the most commonly reported *Salmonella* serotypes, although the SGSA performance was assessed using a panel of *Salmonella* isolates of known serotype representing 55 these serotypes. The remaining four serotypes were unavailable for testing. Overall the correct serotype was assigned for 95.2% of samples tested, with 100% sensitivity with a specificity ≥ 99% for *S.* Enteriditis, *S.* Typhimurium and a panel of seven other clinically important serotypes. However *S.* Enteriditis isolates were not distinguished unequivocally from other serotypes that possessed very similar antigenic formulae. Additional shortcomings relating to data analysis and sample processing must be overcome before the test can be used for routine testing. The SGSA can potentially provide a rapid and accurate test for the identification of the most commonly reported *Salmonella* serotypes.

The Check & Trace array is a SNP based array that is currently marketed and available in kit form from CHECKPOINTS. Within this project a total of 2135 isolates representing 171 serotypes was evaluated. The technology is developed and marketed by Checkpoints in The Netherlands. The company have been highly proactive in training AHVLA staff and giving unlimited access to their expertise. The C&T array generates a surrogate code based on nucleic acid sequence differences (SNPs) and these are correlated to antigenic formulae of the White-Kauffmann-Le Minor scheme. Currently, the C&T array includes the determination of 300+ of the most commonly reported

Salmonella serotypes, although the performance was assessed using a panel of *Salmonella* isolates representing 171 these serotypes. The C&T array has accreditation for 102 serotypes with the AOAC-RI and 22 of those serotypes with current and future regulatory significance also have International OIE certification. Overall the correct serotype was assigned for 97.84% of samples tested, with 100% sensitivity with a specificity ≥ 99% for *S.* Enteritidis, *S.* Typhimurium and the remainder of the top 15 serotypes. The C&T array can provide a rapid and accurate test for the identification for over 300 reported *Salmonella* serotypes and is commercially available for use.

Project Report to Defra

8. As a guide this report should be no longer than 20 sides of A4. This report is to provide Defra with details of the outputs of the research project for internal purposes; to meet the terms of the contract; and to allow Defra to publish details of the outputs to meet Environmental Information Regulation or Freedom of Information obligations. This short report to Defra does not preclude contractors from also seeking to publish a full, formal scientific report/paper in an appropriate scientific or other journal/publication. Indeed, Defra actively encourages such publications as part of the contract terms. The report to Defra should include:
   - the objectives as set out in the contract;
   - the extent to which the objectives set out in the contract have been met;
   - details of methods used and the results obtained, including statistical analysis (if appropriate);
   - a discussion of the results and their reliability;
   - the main implications of the findings;
   - possible future work; and
   - any action resulting from the research (e.g. IP, Knowledge Exchange).

Final report following on from the Interim Report (September 2011) and the Annual Report (November 2010).

Objective 1: Select commercially available DNA microarray assays that fulfil the criteria set by the EU

Objective 2: Perform a comparative analysis between the existing VLA system (Kauffmann White) and the commercially selected microarray(s)

Objective 3: Perform comparative analysis of the chosen commercial microarray and the VLA developed microarray against the current Kauffmann White gold standard (validated to ISO16140 standard)

Objective 4: Adapt the microarray chosen to an automated system

Objective 5: Using the data obtained from objectives 2, 3 and 4 provide full validation up to test code status which also satisfies the ISO 16140 standard.

The aim of the project was to evaluate alternative molecular methods of *Salmonella* detection and to provide evidence of feasibility for a rapid *Salmonella* alerting system. The series of objectives was designed to facilitate this and all have been met in full. A total of four arrays were assessed. The Premitest array was described in full in the November 2010 annual report (a revised updated version from Checkpoints was analysed in full and will be described here [see below]). The Luminex based array was an early stage prototype out of CDC Atlanta. This system did not yield any meaningful and the findings methodologies were described in the interim report of November 2011. This meant that a fully comprehensive comparative analysis between all four arrays was not possible. Once fully developed and validated the array may provide a cheaper automated rapid *Salmonella* alerting system but would require rigorous assessment and validation.

The two further arrays; The *Salmonella* Genotyping Array (SGSA) and the Check & Trace array (C&T) were both assessed and the C&T array validated to AHVLA test code status standard. In addition both arrays are in an format ready for automation although the SGSA can perform 96 samples in one batch compared to the C&T 36 samples. The details of both assessments are now described.

The *Salmonella* genoserotyping array (SGSA)

The SGSA identifies serotypes through its capability to detect the unique sequences which encode the many unique somatic and flagellar antigens described by the White-Kaufmann-Le Minor scheme. The SGSA was developed by three partner laboratories (Public Health Agency of Canada (PHAC), Austrian Institute of Technology (AIT) and the Animal Health and Veterinary Laboratories Agency (AHVLA)) and has been recently described (Franklin *et.al.* 2011).

The SGSA is a rapid test which can determine the serotype of an isolate within two days of obtaining a bacterial culture. This is significantly faster than traditional *Salmonella* serotype methods, which can take four days or more to establish serotype, especially when the identification of second phase flagellar antigens is required for definitive typing. The fast SGSA 'turn-around-time' can enable rapid responses to be mounted in outbreak situations. Up to 96 samples can be processed in parallel by SGSA, providing significant efficiency savings. The SGSA does not use polyclonal antisera and thereby overcomes variations in the quality and availability of these reagents. Furthermore it obviates the significant animal welfare implications arising from the preparation and use of antisera.

The SGSA has been designed to identify 59 *Salmonella* serotypes, representing the most commonly reported *Salmonella* serotypes from human and non-human sources (Table 1).

Table 1. The Salmonella serotypes targeted by the SGSA.

S. Abony	*S.* Goldcoast	*S.* Oranienburg
S. Agona	*S.* Hadar	*S.* Orion
S. Albany	*S.* Heidelberg	*S.* Paratyphi A
S. Amsterdam	*S.* IIIb 61:k:1,5,(7)	*S.* Paratyphi B var. Java
S. Anatum	*S.* Indiana	*S.* Panama
S. Arizonae	*S.* Infantis	*S.* Pullorum
S. Blockley	*S.* Java	*S.* Rissen

S. Bovismorbificans	*S.* Javiana	*S.* Saintpaul
S. Braenderup	*S.* Kedougou	*S.* Sandiego
S. Brandenburg	*S.* Kentucky	*S.* Schwarzengrund
S. Bredeney	*S.* Kiambu	*S.* Senftenberg
S. Cerro	*S.* Kottbus	*S.* Stanley
S. Chester	*S.* Livingstone	*S.* Stanleyville
S. Choleraesuis	*S.* London	*S.* Tennessee
S. Corvallis	*S.* Mbandaka	*S.* Thompson
S. Derby	*S.* Mississippi	*S.* Typhi
S. Dublin	*S.* Montevideo	*S.* Typhimurium
S. Enteritidis	*S.* Muenchen	*S.* Virchow
S. Gallinarum	*S.* Muenster	*S.* Weltevreden
S. Give	*S.* Newport	

This list of *Salmonella* serotypes targeted by the SGSA was carefully compiled from peer reviewed publications and reference laboratory data. It includes the top ranked serotypes from a diversity of sources for human isolates (Table 2) and non-human isolates (Table 3). Detailed rankings for each serotype are presented in Appendices 1 and 2.

Table 2. Most commonly identified Salmonella serotypes reported from human isolates targeted by the SGSA.

Geographical Region	**Year**	**Serotypes targeted by SGSA**	**Reference**
EU	2008	10 most commonly isolated	Anonymous 2010a
EU	2007	10 most commonly isolated	Anonymous 2010a
UK	2008	20 most commonly isolated	Health Protection Agency (UK)
Austria	2008	10 most commonly isolated	Austrian Research Centres (Austria)
USA	2005	13 most commonly isolated	Centers for Disease Control and Prevention (USA)
Canada	2007	15 most commonly isolated	National Microbiology Laboratory (Canada)
Africa	2001 - 2007	14 most commonly isolated	Hendriksen *et.al.* 2011
Asia	2001 - 2007	20 most commonly isolated	Hendriksen *et.al.* 2011
Europe	2001 - 2007	20 most commonly isolated	Hendriksen *et.al.* 2011
N. America	2001 - 2007	20 most commonly isolated	Hendriksen *et.al.* 2011
Oceania	2001 - 2007	6 most commonly isolated	Hendriksen *et.al.* 2011
Latin America	2001 - 2007	8 most commonly isolated	Hendriksen *et.al.* 2011
Global	2000 - 2002	20 most commonly isolated	Galanis *et.al.* 2006
Africa	2000 - 2002	4 most commonly isolated	Galanis *et.al.* 2006
Asia	2000 - 2002	5 most commonly isolated	Galanis *et.al.* 2006
Europe	2000 - 2002	5 most commonly isolated	Galanis *et.al.* 2006
Latin America and Carribean	2000 - 2002	5 most commonly isolated	Galanis *et.al.* 2006
N. America	2000 - 2002	5 most commonly isolated	Galanis *et.al.* 2006
Global	1995	20 most commonly isolated	Herikstad *et.al.* 2002

Table 3. Most commonly identified Salmonella serotypes reported from non-human isolates targeted by the SGSA.

Source	Year	Serotypes targeted by SGSA	Reference
GB: Animal	2007	20 most commonly isolated	Veterinary Laboratories Agency (UK)
Austria: Animal	2008	10 most commonly isolated	Austrian Research Centres (Austria)
Canada: Animal	2008	20 most commonly isolated	Laboratory for Foodborne Zoonoses (Canada)
EU: Broiler meat	2008	10 most commonly isolated	Anonymous 2010a
EU: *Gallus gallus* flocks	2008	10 most commonly isolated	Anonymous 2010a
EU: Pig meat	2008	10 most commonly isolated	Anonymous 2010a
EU: Pig herds	2008	10 most commonly isolated	Anonymous 2010a
EU: Bovine meat	2008	10 most commonly isolated	Anonymous 2010a
EU: Cattle herds	2008	10 most commonly isolated	Anonymous 2010a
Global: non-human	2000 - 2002	9 most commonly isolated	Galanis *et.al.* 2006
Africa: non-human	2000 - 2002	5 most commonly isolated	Galanis *et.al.* 2006
Asia: non-human	2000 - 2002	6 most commonly isolated	Galanis *et.al.* 2006
Europe: non-human	2000 - 2002	5 most commonly isolated	Galanis *et.al.* 2006
Latin America and Carribean: non-human	2000 - 2002	5 most commonly isolated	Galanis *et.al.* 2006
North America: non-human	2000 - 2002	5 most commonly isolated	Galanis *et.al.* 2006

The performance of the SGSA was assessed at the AHVLA by testing 335 *Salmonella* isolates and comparing results to known serotype (as established by classical serotyping methodology). The repeatability and analytical specificity were also assessed using well characterised strains. This report details the results of this assessment.

Similar studies are being undertaken at the PHAC and AIT. When these studies have been completed all data will be compiled and SGSA performance at the three laboratories assessed.

Methods

Salmonella Isolates

Isolates were obtained from the collections held by the AHVLA *Salmonella* Reference Laboratory. The serotype of each isolate had previously been determined using traditional serological methods according to established procedures (AHVLA SOP CBU0001). Isolates from the reference collection have been serotyped many times however the field isolates were only serotyped once by the KW system. *S.* Arizonae, *S.* Java, *S.* Paratyphi A, and *S.* Typhi were not tested as no isolates were available in the AHVLA reference collections.

DNA Extraction

Bacteria were cultivated overnight at 37°C on Blood Agar Base No. 2+ Lactose + Neutral Red or 5% Sheep Blood Agar plates. A 1 µl loopful of cells was recovered from the plate and resuspended in 180 µl of Buffer ATL (Qiagen, Crawley, UK). Following addition of 20 µl of Proteinase K (20mg/ml) and 10 µl of lysozyme (10mg/ml) the sample was mixed thoroughly by vortexing and incubated at 59°C for between 1 hour to 3 hours. After the addition of 4 µl of Qiagen RNaseA solution (100 mg/ml) the DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturuer's protocol. Eluted DNA was used immediately or stored at -20°C.

SGSA Method

The SGSA method was performed essentially as described in Franklin *et.al.* (2011). In brief, DNA extracts were PCR amplified in three separate multiplex reactions. Aliquots from the three reactions were pooled and treated with shrimp alkaline phosphatase (SAP) to dephosphorylate the remaining nucleotides. The SAP-treated PCR products were then biotin labelled by the sequence specific end labelling of oligonucleotides (SSELO) method. Samples were hybridized to the SGSA using a hybridization kit (Alere Technologies, Jena,

Germany) and processed according to the manufacturer's instructions, except the hybridization was carried out for 1 h at 60°C and bound with D1 substrate reagent for 15 min at room temperature. Signal intensities were detected using an ArrayMate reader (Alere Technologies).

SGSA Data Analysis

The signal intensity results were analysed using a Microsoft Excel macro, in which each *Salmonella* sample was characterized by a unique probe pattern based on the identification of an O antigen and of phase 1 and phase 2 flagellar antigens. The antigenic formula was then used to designate serotype according to the White-Kauffmann-Le Minor scheme (Grimont and Weill, 2007).

In some instances, the SGSA is unable to differentiate between the serotypes of isolates that have very similar antigenic formulae. For example, *S.* Enteritidis is identified by the SGSA as: *"S.* Enteritidis (1,9,12:g,m), *S.* Blegdam (9,12:g,m,q:–), *S.* Moscow (1,9,12:g,q:–), *S.* Rosenberg (9,12:g,z_{85}:–), *S.* Hillingdon (9,46:g,m:–), *S.* II (1,9,12:g,z62:[e,n,x]), or *S.* Wuppertal (9,46:z_{41}:–)". In this evaluation, when the SGSA designated more than one possible serotype, all serotypes were reported, and the result was considered correct if any of the designated serotypes matched the serotype obtained by the classical serotyping method.

Monophasic variants of *S.* Typhimurium-like strains, lacking the *fljB*-encoded second phase H antigen (*i.e.* with the antigenic structure 1,4,[5],12:i:-) are of increasing importance. The current EFSA scientific opinion (Anonymous 2010b) states that monophasic strains with the basic antigenic formula 1,4,[5],12:i:- are to be regarded as variants deriving from *S.* Typhimurium, and reported as 'monophasic *S.* Typhimurium'. The SGSA reports the antigenic formula 1,4,[5],12:i:- as B:i:- and designates the sample as 'possible monophasic *S.* Typhimurium'. Supplementary testing to establish this conclusively would be required. Some monophasic *S.* Typhimurium isolates possess silent second phase genes. These would be detected by the SGSA and reported as *S.* Typhimurium, not as monophasic *S.* Typhimurium. This is a correct result as monophasic *S.* Typhimurium are variants deriving from *S.* Typhimurium.

The serotype determined by SGSA was compared to the known serotype. Samples which did not give the expected result were repeat tested to ensure the error was not a consequence of a technical error (*e.g.* sample contamination).

The SGSA has the theoretical capability to identify 1,358 of the 2,579 *Salmonella* serotypes described in the White-Kaufmann-Le Minor scheme including 1026 of the 1531 *S. enterica salmonella* serotypes. The aim of this evaluation was to assess the capability of the SGSA to detect the 59 most commonly reported serotypes (Table 1). Therefore SGSA serotype designations that were not from these 59 serotypes were scored as 'not identified'.

Results

Testing of *Salmonella* Isolates

A panel of 216 *Salmonella* isolates representing 55 of the 59 serotypes listed in Table 1 was assembled from the AHVLA collections. Isolates from the serotypes *S.* Arizonae, *S.* Java, *S.* Paratyphi A, and *S.* Typhi are not held in the AHVLA collections and were therefore not tested.

The *Salmonella* isolates were tested by SGSA and repeats performed as required (30 isolates needed repeating). The SGSA correctly identified the serotype for 207 / 216 (95.8%) samples tested (Table 4). Of the 55 serotypes examined 49 were consistently identified by SGSA, including *S.* Enteriditis and *S.* Typhimurium. Three variant forms of *S.* Typhimurium were tested and have been analysed together in Table 4. The 14 *S.* Typhimurium and 6 *S.* Typhimurium var. Copenhagen tested were all designated *S.* Typhimurium by SGSA. Of the seven monophasic *S.* Typhimurium isolates tested by SGSA, five were designated as 'possible monophasic *S.* Typhimurium' and two as *S.* Typhimurium.

For some isolates of particular serotypes, the correct probes were positive on the microarray but were erroneously analysed. The reporting system uses a macro in Excel macro that in early versions was erroneous, therefore, identifying incorrect serotypes. On detailed analysis and correction of the macro, all these were scored as correctly identified in the final analysis presented in Table 4 (column 6).

Table 4. Salmonella serotype designations determined by SGSA for panel of reference isolates.

Serotype	Total Number of Isolates Tested	Number of Isolates Correctly Identified	Number of Isolates Repeat Tested by SGSA	Incorrect serotype designation	Serotypes with Macro Analysis Error
S. Abony	3	3	2		
S. Agama	2	2	0		
S. Agona	3	3	2		
S. Albany	1	1	1		
S. Amsterdam	2	2	1		
S. Anatum	3	3	1		
S. Arizonae	not tested	not tested	not tested		
S. Blockley	3	0	3	3 not identified	
S. Bovismorbificans	3	1	2	2 not identified	
S. Braenderup	3	3	0		
S. Brandenburg	1	1	0		
S. Bredeney	3	3	0		1 Macro error
S. Cerro	3	3	1		
S. Chester	1	1	0		
S. Choleraesuis	1	1	0		
S. Corvallis	3	3	1		
S. Derby	3	3	2		
S. Dublin	3	3	1		
S. Enteritidis	20	20	0		
S. Gallinarum	4	4	N/A		4 Macro errors
S. Give	3	3	N/A		3 Macro errors
S. Goldcoast	2	2	0		
S. Hadar	7	7	0		
S. Heidelberg	7	7	0		
S. IIIb 61:k:1,5,(7)	3	3	0		
S. Indiana	3	3	1		1 Macro error
S. Infantis	7	7	0		
S. Java	not tested	not tested	not tested		
S. Javiana	3	3	0		
S. Kedougou	3	3	0		
S. Kentucky	7	7	0		1 Macro error
S. Kiambu	3	3	0		1 Macro error
S. Kottbus	3	3	0		
S. Livingstone	3	3	0		
S. London	1	1	0		
S. Mbandaka	3	3	0		
S. Mississippi	3	3	0		
S. Montevideo	7	7	0		3 Macro errors
S. Muenchen	3	3	1		
S. Muenster	1	1	0		
S. Newport	7	7	2		
S. Oranienburg	2	2	0		1 Macro error
S. Orion (Binza &Thomasville)	2	2	0		1 Macro error
S. Paratyphi A	not tested	not tested	not tested		
S. Paratyphi B var. Java	3	3	0		
S. Panama	1	1	N/A		1 Macro error
S. Pullorum	3	3	N/A		3 Macro errors
S. Rissen	3	2	1	*1 designated S.* Typhimurium	
S. Saintpaul	3	3	0		
S. Sandiego	1	1	0		

S. Schwarzengrund	3	3			3 Macro errors
S. Senftenberg	3	3	2		
S. Stanley	3	2	1	*1 designated S.* Muenchen	
S. Stanleyville	3	3	0		
S. Tennessee	3	2	3	1 not identified	
S. Thompson	3	3	0		
S. Typhi	not tested	not tested	not tested		
S. Typhimurium (inc. monopahsic and var. Copenhagen)	27	27	1		
S. Virchow	7	7	0		
S. Weltevreden	2	1	1	1 not identified	
Total	**216**	**207**	**30**	**9**	**23**

Not tested = serotype not tested at AHVLA; N/A = not applicable

After correction of the Excel macro, a total of nine isolates from six serotypes were given an incorrect serotype designation by SGSA (Table 4). Seven isolates were designated as 'not identified', and two were misidentified as a different commonly identified serotype (Table 4). The serotypes that could not be consistently identified were *S.* Blockley, *S.* Bovismorbificans, *S.* Rissen, *S.* Stanley, *S.* Tennessee, and *S.* Weltevreden.

To further examine SGSA specificity, an additional 119 isolates were tested. These isolates represent rare serotypes that the SGSA was not designed to identify (*i.e.* non-target serotypes). For these samples 112 / 119 (94.1%) were correctly designated as 'not identified', *i.e.* they were not identified as a serotype from Table 1 (data not shown). Three samples were misidentified as a target serotype due to a macro analysis error but were scored as correctly identified in this analysis because an update to the macro will resolve these software errors (Table 5). The seven non-target serotypes that were incorrectly identified as a target serotype are listed in Table 5. For three of these samples the misidentification arose due to the absence of a probe for one of the antigens present in the tested sample. For the rough:g,m:- isolate a somatic antigen type had not been assigned by traditional serotyping, while the SGSA detected the gene encoding the O:9 (D1) antigen and the sample was consequently designated *S.* Enteritidis. The presence of this gene will need to be verified by PCR and sequencing to determine if the SGSA has correctly identified the sample. Four samples were incorrectly designated as a target serotype (Table 5).

Table 5. Incorrect Salmonella serotype designations determined by SGSA from a panel of non-target reference isolates.

Serotype Tested	**SGSA designation**	**Explanation of incorrect designation**
S. Kiel	*S.* Dublin	Macro analysis error (scored as correct designation)
S. Panama	*S.* Gallinarum	Macro analysis error (scored as correct designation)
S. Rostock	*S.* Enteritidis	Macro analysis error (scored as correct designation)
S. Aesch	*S.* Newport	Absence of probe for z60 flagellar antigen
S. Champaign	*S.* Thompson	Absence of probe for O:39 (Q) somatic antigen
S. Crossness	*S.* Heidelberg	Absence of probe for 0:67 somatic antigen
ROUGH:g,m:-	*S.* Enteritidis	PCR for antigen O:9 (D1) gene required
S. Aarhus	*S.* Cerro	Incorrect serotype designation
S. Altona	*S.* Bovismorbificans	Incorrect serotype designation
S. Menston	*S.* Montevideo	Incorrect serotype designation
S. Naestved	*S.* Dublin	Incorrect serotype designation

SGSA Sensitivity and Specificity

The serotype designations obtained with the SGSA were compared to the true serotype (as determined by traditional serotyping) and assay performance assessed. SGSA sensitivity and specificity were determined by 2 x 2 box analysis for *S.* Enteritidis, *S.* Typhimurium, and a group of seven other common serotypes (*S.* Hadar, *S.* Heidelberg, *S.* Infantis, *S.* Kentucky, *S.* Montevideo, *S.* Newport, and *S.* Virchow), Table 6.

Table 6. SGSA sensitivity and specificity.

	***S.* Enteritidis**	***S.* Typhimurium**	**Other common serotypes [b]**
True Positive	20	27	50
True Negative	314	307	282
False Positive	1 [a]	1	3
False Negative	0	0	0
Sensitivity	**100%** (Lower 95% CI: 83.2%)	**100%** (Lower 95% CI: 87.2%)	**100%** (Lower 95% CI: 92.89%)
Specificity	**99.7%** [a] (95% CI: 98.2% – 99.9%)	**99.7%** (95% CI: 98.2% – 99.9%)	**99.0%** (95% CI: 97.0% – 99.8%)

[a] The Rough:g,m:- isolate has been scored as a False Positive for this analysis.
[b] *S.* Hadar, *S.* Heidelberg, *S.* Infantis, *S.* Kentucky, *S.* Montevideo, *S.* Newport, *and S.* Virchow.

SGSA analytical specificity

DNA extracts from five non-*Salmonella* bacterial genera were tested by SGSA to assess analytical specificity. The number of isolates tested and SGSA results are given in Table 7. The majority of samples (16 / 18) were not assigned a *Salmonella* serotype and therefore the SGSA performed correctly.
However, one *Escherichia coli* sample was incorrectly designated *S.* Typhimurium and another was designated *S.* Virchow.

Table 7. SGSA testing of non-Salmonella isolates.

Sample	**Number of Isolates Tested**	**SGSA designation**
Citrobacter spp.	3	3 not identified
Escherichia coli	6	4 not identified; 2 assigned *Salmonella* serotype
Klebsiella spp.	3	3 not identified
Proteus spp.	3	3 not identified
Staphylococcus aureus	3	3 not identified

SGSA Repeatability

The repeatability of the SGSA was assessed by the blind testing of five *Salmonella* samples in triplicate on four separate occasions. The samples used were those from the *Salmonella* Serotyping Proficiency Test 0084, prepared and provided by the AHVLA Quality Assurance Unit. On each occasion one blank (water) negative control test was also performed. The SGSA correctly identified the serotype 60 / 64 (93.8%) times (the water controls being designated as 'not identified'), Table 8. For the *S.* Dublin isolates the correct probes were positive on the microarray but were sometimes erroneously analysed by the Excel macro and therefore identified as an incorrect serotype. These were scored as correctly identified in the final analysis presented in Table 8, as an update to the macro will resolve these software errors. Investigation of the four misidentifications suggest that they resulted from failed or sub-optimal PCRs.

Table 8. Repeatability testing of SGSA.

Sample	**Correct identification / Total tested**	**Notes**
Sample A (*S.* Dublin)	12 / 12	macro analysis error
Sample B (*S.* Agama)	12 / 12	
Sample C (*S.* Typhimurium)	12 / 12	
Sample D (*S.* Goldcoast)	11 / 12	Possible failed PCR
Sample E (*S.* Infantis)	9 / 12	Possible sub-optimal PCR
Water	4 / 4	
Final Correct / Total	60 / 64	

Discussion

The SGSA has been designed to identify the 59 most commonly reported *Salmonella* serotypes from human and non-human sources. When tested with a panel of 335 isolates (that included 216 isolates from 55 of the common serotypes) the SGSA gave the correct serotype designations for 319 samples (95.2%). It correctly identified 207 / 216 (95.8%) of the 55 most common serotypes tested. The SGSA is able to reliably identify most (49 / 55) of the commonly identified *Salmonella* serotypes tested, but six target serotypes could not be consistently identified (*S.* Blockley, *S.* Bovismorbificans, *S.* Rissen, *S.* Stanley, *S.* Tennessee, and *S.* Weltevreden).

The capability of the SGSA to correctly identify the four serotypes *S.* Arizonae, *S.* Java, *S.* Paratyphi A, and *S.* Typhi was not examined. This is because these serotypes are not held in the AHVLA *Salmonella* collections. The partner laboratories at PHAC and AIT are undertaking parallel evaluations of the SGSA. They have assembled and are testing panels of *Salmonella* isolates from their collections (including *S.* Arizonae, *S.* Java, *S.* Paratyphi A, and *S.* Typhi). Therefore the SGSA will be assessed for these serotypes in due course.

Test sensitivity was 100% for *S.* Enteriditis and *S.* Typhimurium. This indicates that there were no false negative results for these clinically and legislatively important serotypes, all having been correctly identified by SGSA. Seven isolates identified as monophasic *S.* Typhimurium by traditional serotyping were tested by SGSA. Two isolates were positive for the second phase flagellar antigens and designated as *S.* Typhimurium. The genes encoding these antigens are likely to be present but silent (*i.e.* not expressed) in these isolates. This is the correct identification as monophasic *S.* Typhimurium are variants deriving from *S.* Typhimurium (Anonymous 2010b). The SGSA correctly identified the 20 *S.* Enteriditis isolates tested. However, a limitation of the assay is that it cannot unequivocally distinguish these from other serotypes possessing very similar antigenic formulae. These other serotypes are reported at much lower frequencies than *S.* Enteriditis, but this still represents a significant limitation in the capability of the assay. Unequivocal identification of *S.* Enteriditis should be introduced in subsequent iterations of the SGSA, possibly by use of serotype specific PCR and/or hybridisation probes.

The test specificity was 99.7% for *S.* Enteriditis and *S.* Typhimurium. One isolate (*S.* Rissen) was false positive by SGSA for *S.* Typhimurium, and this sample will be further tested to establish the reason for the error. One *S.* Rostock sample was misidentified as *S.* Enteriditis as a result of a macro analysis error (and therefore not considered a false positive). The rough:g,m:- isolate was identified as *S.* Enteriditis by SGSA, due to the detection of the gene encoding the D1 somatic antigen. This gene may be present in this isolate but not expressed and therefore not detected by classical serotyping methods. PCR and sequencing of amplicon will be required to verify the presence of the gene in this isolate. There iis also the issue of the SGSA being unable to differentiate between the serotypes of isolates that have very similar antigenic formulae e.g. *S.* Enteritidis is identified by the SGSA as: *"S.* Enteritidis (1,9,12:g,m), *S.* Blegdam (9,12:g,m,q:–), *S.* Moscow (1,9,12:g,q:–), *S.* Rosenberg (9,12:g,z_{85}:–), *S.* Hillingdon (9,46:g,m:–), *S.* II (1,9,12:g,z62:[e,n,x]), or *S.* Wuppertal (9,46:z_{41}:–). This is an issue that would need refinement before use as a robust rapid alerting system.

A small but significant fraction of the isolates from the common serotypes needed retesting due to technical errors (30 / 216; 13.9%). These samples were retested because investigation of the results indicated that there was likely to be cross-contamination of samples. Upon retesting, many samples were correctly identified. The cross-contamination of samples is likely to have arisen when the PCR products were pooled for use in the SAP reaction. The SGSA has good repeatability with 60 / 64 tests returning the correct result. The incorrect results appear to have arisen as a consequence of failed or poor PCR reactions. Improvements to the handling procedures need to be implemented to address the susceptibility of this test to cross-contamination and the potential for impaired PCR reactions. It should be borne in mind that this system is new and that acquiring the skill of the method did take time and practice. Error rates are likely to drop significantly with improved familiarity with using the SGSA system.

The parallel evaluations being undertaken by the partner laboratories will further inform on SGSA performance. Each partner is testing *Salmonella* isolates from their North American (PHAC) or continental European (AIT) collections. Thus a wide diversity of samples from different geographical regions will be tested. These data will be combined with that presented in this report to further characterise the specificity, sensitivity, and accuracy of the SGSA. Repeatability studies (using the same *Salmonella* Serotyping Proficiency Test samples employed in this work) will also be undertaken by each partner, enabling assessment of test repeatability and reproducibility.

The SGSA signal intensity results were analysed using a Microsoft Excel macro and serotype assigned based on the unique probe pattern obtained. A number of errors in the macro were uncovered during the course of this study (*e.g.* 23 / 216 in the panel of common serotypes). These errors resulted in the incorrect serotype designation when the correct probes were positive on the microarray. In the analysis undertaken here it has been assumed that these errors can be fixed by making improvements to the macro, and that reanalysis of the results with the improved macro would result in the correct serotype designation. Without a correctly functioning analysis macro the performance of the SGSA will be seriously eroded.

Two *E. coli* isolates were identified as *Salmonella* serotypes, representing a potential source of error. To reduce or prevent such errors procedures must be in place to ensure that only *Salmonella* isolates are tested by SGSA. This can entail correct identification of *Salmonella* following established procedures (e.g. AHVLA SOP BA.198) and the separation of the handling of *Salmonella* from other bacteria.

If the shortcomings identified in this assessment can be overcome the SGSA could be a rapid and accurate test for the identification of most of the commonly reported *Salmonella* serotypes. Used as an initial serotyping test, it could rapidly identify the serotypes for the majority of *Salmonella* isolates submitted to testing laboratories. Samples not identified would then need to be tested by traditional serotyping methods. With further validation and the ability to identify many serotypes beyond those tested in this study, the SGSA has the potential to be a very powerful tool.

The Check & Trace SNP based array (C&T)

The C&T array confirms the presence and serotype of *Salmonella* from a pure culture with a single test and in one day, which significantly decreases the serotyping lead times enabling a quick tracing and therefore elimination of the source. The equipment required can be found in most molecular laboratories and there is also a two day training course for those with no experience in molecular techniques. The analysis uses automated software based interpretation of results. The system uses a SNP based DNA technology and so updating the C&T system to include additional serotypes is by updating the software rather than re-validating the whole array.

The C&T discriminates *Salmonella* serotypes through differences in their DNA sequence. Even non-typable isolates can be assigned to a serotype, thereby offering a greater discriminative power than conventional techniques. The technology uses a microarray platform. Each position on the microarray represents a specific DNA marker associated with a unique *Salmonella* target sequence; and they only become visible if the DNA markers exactly match the corresponding DNA sequences of the *Salmonella* isolate. The combination of present and absent markers yields a pattern.

The C&T array currently recognises and identifies over 300 serotypes and has AOAC-RI certified for 102 serotypes; International OIE certification for the 22 serotypes of current and future regulatory significance; has growing number of serotypes detected supported by a highly skilled team at Checkpoints. Again the use of a molecular system means that the C&T array does not use polyclonal antisera and thereby overcomes variations in the quality and availability of these reagents and obviates the significant animal welfare implications arising from the preparation and use of antisera. The array has recently undergone automation and can serotype up to 36 samples on one run. The Checkpoints team are currently working on increasing this number significantly.

Table 9 showing the list of AOAC-RI certified Salmonella serotypes using the C&T array (bold is also OIE certified)

1,4,[5],12:i:-	Aberdeen	Abony	Adelaide	**Agona**
Albany	Altona	**Anatum**	Banana	Bareilly
Berta	Blockley	Bongori	Bovismorbificans	Braenderup
Brandenburg	**Bredeney**	Carrau	Cerro	Chandans
Chester	Choleraesuis	Coeln	Colindale	Corvallis
Cubana	**Derby**	**Dublin**	Duisburg	Eboko
Enteritidis	Gallinarum Gallinarum	Gallinarum Pullorum	Give	Gloucester
Goldcoast	Grumpensis	**Hadar**	Havana	**Heidelberg**
Ibadan	Idikan	**Indiana**	**Infantis**	Isangi
Jangwani	Javiana	Kedougou	Kentucky	**Kottbus**

Lexington	Lille	Litchfield	Liverpool	Livingstone
London	Manchester	Manhattan	Matadi	**Mbandaka**
Meleagridis	Michigan	Mikawasima	Minnesota	Monschaui
Montevideo	Muenchen	Muenster	Napoli	**Newport**
Ohio	Oranienburg	Orion	Orion	Oslo
Ouakam	Panama	Paratyphi A	**Paratyphi B**	Paratyphi B v Java
Paratyphi C	Pomona	Poona	Regent	Rissen
Rubislaw	**Saintpaul**	Sandiego	Schwarzengrund	**Senftenberg**
Stanley	Stourbridge	Telelkebir	**Tennessee**	Thompson
Typhi	**Typhimurium**	Urbana	**Virchow**	Wandsworth
Weltevreden	Worthington	Yoruba		

Salmonella Isolates selected

A comprehensive panel of 2135 *Salmonella* isolates of known serotype were obtained from the culture collection held at AHVLA. 2015 were field isolates collected primarily over 2011. The panel also included 24 non-*Salmonella* isolates (*Escherichia coli*, *Staphylococcus epidermidis* and *Staphylococcus aureus*) as negative *Salmonella* controls. The selected strains included isolates from routine submissions, non *Salmonella* strains and isolates from the validation panel of 104 strains as described above.

All 2135 isolates have been characterised using the classic Kauffmann White (KW) typing scheme at AHVLA. This set of isolates covers 171 different serotypes which include; *Salmonella* Enteriditis, *Salmonella* Typhimurium, *Salmonella* Hadar, *Salmonella* Heidelberg, *Salmonella* Infantis, *Salmonella* Newport, *Salmonella* Montevideo, *Salmonella* Typhimurium 1,4,5,12:i: and *Salmonella* Virchow.

Microarray Layout

The C&TS array utilises 32 hybridisation spots (Figure 1). Each microarray is contained in an array tube (AT) and is able to detect three distinct samples on one array (Set 1, 2 and 3). Each set is equipped with the appropriate controls; *Salmonella* specific probes (spots 13, 14 and 15), negative control (spots 16 and 29), internal control (spots 30, 31 and 32), hybridisation control spots. The array uses one or two available probe sets to determine the serotype, primarily the first set of probes (spots 1-12) are used and where necessary a secondary set (spots 19-26) are taken into account by the software should there be any uncertainties with the first set. Where only set one is used a numerical score is obtained, where the secondary set has been used the genovar score will be an alphanumerical code (eg, *S.* Kedougou 12327.EF).

Figure 1 – DNA microarray layout

		1 (set 1)	2	3	4	5	6	7	8	9	
	10	11	12	13	14	15	16	17	18	19/A	
	20/B	21/C	22/D	23/E	24/F	25/G	26/H	27	28	29	
	30	31	32		1 (set 2)	2	3	4	5	6	
	7	8	9	10	11	12	13	14	15	16	
		17	18	19/A	20/B	21/C	22/D	23/E	24/F	25/G	
	26/H	27	28	29	30	31	32	1 (set 3)	2	3	
	4	5	6	7	8	9	10	11	12	13	
	14	15	16		17	18	19/A	20/B	21/C	22/D	
	23/E	24/F	25/G	26/H	27	28	29	30	31	32	

The test is carried out in 6 stages: (reagents are provided with the C&T kit)

1. Sampling
2. Lysis to obtain DNA
3. DNA recognition step A
4. DNA recognition step B
5. DNA recognition step C
6. Detection step

In brief, the test uses specific DNA probes that generate a collection of ligated probes from a crude genomic DNA lysate. These ligated probes are further amplified by PCR and are subsequently hybridised to DNA microarray spotted with probe-specific oligonucleotides which sits in an array tube (AT). A biotin label is incorporated into the amplified products which is visualised by colorimetric detection. Array images are analysed by a photometric detector. The biotinylated products translate into a unique code, which corresponds to a specific position on the microarray; combination of these codes yields a genovar score which in turn translates to a particular serotype. The C&T software automatically identifies the genovar score and serotype. Results are analysed instantly, so there is no need for time-consuming interpretation. , eg, presumptive *Salmonella* with a genovar score of 50 is identified as *Salmonella* Derby.

C&T array software Assignments

- *Salmonella serotype (genovar score)* – reports the *Salmonella* serotype with (gen sc) e.g. *Salmonella* Derby (50)
- *Salmonella, genovar* – indicates the sample is positive for *Salmonella*, but the unique serovar array pattern i.e. gen sc is not defined in the software. Meaning the serotype most likely does not match one of the serotypes outlined in Appendix 1. However CP has an extensive database containing results for over 300 different serotypes, which is continuously growing. It is possible that the gen sc is present in the CP database (C&T DATABASE) but has not been added to the software. Samples that display a gen sc, the images is sent to serovar@check-points.com where it is compared to the database and feedback sent within two working days. In order for codes to be added to the software there is a requirement by the developers to have a minimum of 3 strains that present the same gen sc and have been confirmed by classical Kauffmann White typing. For example, in the AHVLA data set there were a several gen sc that were routinely observed and could potentially be added to the software e.g. 23 strains were scored as *Salmonella*, genovar 10763 which were all typed as London by the KW scheme.
- *Salmonella suspected, please repeat from the sampling step* – the software did not find sufficient spots to give conclusive result. Inspect the array tube and image, an air bubbles or dust particles may interfere with the result. If this is the case gently tap the AT or gently pipette the liquid up and re-take

the image. If the result remains the same, repeat the test from the sampling step.

- *Salmonella spp. (Serovar cannot be determined*) – the sample is positive for *Salmonella* but the software cannot determine a clear pattern and is therefore unable to specify the serotype. Due to, an air bubble or dust which can be rectified by gently tapping the AT and re-taking the image. Or it could DNA preparation is contaminated inhibiting the reactions in which case the sample would need to be re-made. Or over/under staining of the image. In some cases it is possible to manually determine the serotype by adjusting the image using the software. This should be done with care or sent to CP as incorrect handling could lead to a false positive result.
- *Salmonella (Possibly non enterica subspecies)* – software is unable to provide a serotype as gen sc is given e.g score 3076 is a *S*.enterica subsp houtenae
- *No Salmonella* – software is unable to detect *Salmonella* specific probes

Potential reaction failures:

- *DNA recognition not OK, please reprocess sample from sampling step* – software could not find reaction control spots on the AT. Controls are used to check the performance of steps A and C. Failure can indicate reaction mixtures not prepared properly, inhibiton of the reaction due to not enough material or too much cell debris. Cross contamination, too many spots on the array. In all cases repeat from the sampling step.
- *Hybridisation spots not found: please reprocess sample from step A* – the software did not find the hybridisation spots on the AT. The hybridisation control is used to check if hybridisation of the PC product to the AT has been performed correctly.
- *Reference spots not found: check for air bubbles, if not redo detection step* – due to air bubbles or dust gently remove by tapping AT. Or maybe due to conjugate not removed properly, staining solution not added or spots missing due to damage to the array.
- *Human Error* Where clear from the data that a sample mix up has occurred. This could have been during plating out, DNA preparation, addition of DNA to array tube. Evident from incorrect serotype, where amplification controls present and a clear cut result is obtained. Where there are discrepancies it is advised that these samples be re-serotyped.
- *Overnight Ligation.* Overnight ligation was required to be able to meet target of >2000 strains. Overnight ligation is not typically carried out as it has shown to result in increased background. It was found that spot 11 on the array was enhanced. However, when compared with the overall intensity of the array image spot 11 was seen to be obviously weaker allowing for manual override of the spot and the correct serotype result.

<u>*Results*</u>

Comparative analysis of the sensitivity and specificity of the C&T microarray against the Kauffmann-White (KW) system was performed using a panel of 2135 *Salmonella* isolates of known serotypes from the AHVLA culture collection and routine field isolates. The overall results for all 2135 isolates (171 serotypes) on the first run of testing only using the C&T array showed a 93.44% complete match to the KW result. The remaining 6.55% were due to software identified reaction failures, genovar results shown rather than serotype; or identified human error.
Full raw and cleaned data are presented in the attached appendix 3.

Of the 2135 isolates,

- 2.15% of the inaccuracies were due to human error identified from the C&T reporting.
- 2.34% of inaccuracies were due to reaction failures, which were flagged up by the software and samples would be repeated. It has been shown that on repeating tests from the sampling step, reaction failures have been cleared and a positive result obtained.

The results when retesting was performed which then encompassed 2039 isolates and increased the completed match to 97.84%; This includes; 9.96% which gave a genovar result only; 3.29% which proposed two or more closely related serotypes and 2.06% proposed acceptable related structures.

Table 10 Overview of results with the known human errors and reaction failures removed

Check&Trace *Salmonella* Result	Number	%
Complete Match with KW serotyping	2039	97.84%
Of those 97.84% the following applies;		
Salmonella Genovar result only	203	9.96%
More than one species *Salmonella* proposed or related species proposed (includes overlaps)	67	3.29%
Partially correct structures proposed (excludes overlaps)	42	2.06%
Positive *Salmonella* not detected	0	0%

Analysis of Results

Complete Match with KW serotyping – 97.84%

- A total of 97.84% of the C&T results were a correct match
- The overall result of the comparison between C&T and KW is 97.84% which falls beneath the acceptable EU limits of 99.9% for S. Enteritidis and S. Typhimurium. However, it should be noted that this includes all the 171 serotypes used to asses C&T, including other structures, which could have resulted from partial typing by the classic KW scheme. It also includes strains that could have potentially been mistyped by the KW scheme, examples are

    - S. Altona strain was typed as S. Poona by C&T. Poona is a well characterised C&T type it may be advisable to re-type such strains.
    - S. Apapa strains which were typed as S. Monschaui by C&T, on comparison of the KW there appears to be failure with the O antigen. Where the S. Monschaui O anitgen = 35 and S. Apapa = 45, it is likely that there has been a failure with classic typing or a typo on entering results.

- When the top 15 serotypes (including *S*. Enteritidis and *S*. Typhimurium, and the vaccine strains) are analysed separately (excluding known human errors and reaction failures) a 100% match was obtained. This exceeds the EU regulation guidelines for all the serotypes listed. (Table 11)

Table 11 – C&T analysis on top 15 serotypes (excluding known errors) using C&T array

Top 15 Serotypes	**Number of samples screened**	**Complete Match**	**Human Error**	**Reaction Failure**	**Overlap**	**genovar core unique to AHVLA samples**	**Overall Match***
S. Enteritidis	53	53	0	0	0	0	100%
S. Dublin	174	167	5	2	0	0	100%
S. Typhimurium	324	317	0	7	0	0	100%
S. Derby	239	233	2	4	0	0	100%
S. Kedougou	80	76	4	0	0	0	100%
S. Virchow	43	39	1	3	0	0	100%
S. Mbandaka	102	99	2	1	0	0	100%
4,5,12:i:-	296	280	7	9	0	0	100%
S. Java	1	1	0	0	0	0	100%
S. Senftenberg	41	40	1	0	0	0	100%
S. Montevideo	89	84	4	1	0	0	100%
S. Tennessee	9	9	0	0	0	0	100%
S. Agona	9	8	0	1	0	0	100%
S. Havana	6	5	0	0	0	1	100%
S. Ohio	7	7	0	0	0	0	100%
Sub-total	1473	1418	26	28	0	1	
TOTAL (with known errors removed)	**1419**						

Salmonella Genovar result only – 9.96%

- 9.96% gave *Salmonella* genovar only. This included samples; that are not recognised by the C&T software and C&T DATABASE, strains that are *Salmonella* non enterica spp., scores that were found to be unique to the AHVLA data set and could be added to the C&T software. It does however still recognise that a *Salmonella* species is present due to the DNA recognition of the O and H antigens.

a) Genovar score 10760 – S.Agama, code specific to AHVLA data set, code seen throughout data set for Agama
b) Genovar score 6176 – S. Derby, code specific to AHVLA data set, code seen throughout data set for S. Derby.
c) Genovar score 12840 – S. Kisarawe, code specific to AHVLA data set, code seen throughout data set for S. Kisarawe.
d) Genovar score 13063 – S. London, code specific to AHVLA data set, code seen throughout data set for S. London.

More than one species Salmonella proposed or related species proposed (includes overlaps) – 3.29%

- 3.29% of the tested samples gave a score that did not match the proposed serotype, these were due to operator error or overlap between serotypes. The isolates that did not match due to human error were strains that have been well characterised both within the C&T DATABASE and within the AHVLA panel of strains, before being able to declare these as incorrect it is strongly advised that these strains are re-serotyped using the classic KW scheme as the mismatches are between samples that are serotypically different. Without re-serotyping the strains it is not possible to say whether there is an issue with the C&T method or an error has been made with the classic KW scheme

  With respect to overlap, there are a known number of overlaps that have been seen with C&T, resulting from similarities between KW structures. These were determined to be partially correct, but not accurate enough to be a reportable result. This also includes where two or more serotypes have been assigned by C&T, the examples seen have included;

1. *Salmonella* Binza (structure 3,15: y : 1,5) is a known variant of S. Orion (structure 3,10: y : 1,5) as determined by KW. It would be considered an acceptable result using a microarray but noted that KW can determine between the two.

2. *Salmonella* California and *Salmonella* Banana are totally identical in structure; the difference between the two serotypes is determined by a biochemical reaction with inositol. The C&TS array can currently identify Banana but not California although this would also be considered an acceptable result.

3. *Salmonella* Typhimurium, typed as Typhimurium or 1,4,[5],12:i:- The C&T test does present some known limitations (described in full on page 22 of the user manual), with the typing of S. Typhimurium and monophasic S. Typhimurium, where 1,4,[5],12:i:- can be typed as S. Typhimurium and vice a versa. There is only one spot difference on the array between two samples relating to the FljA gene, when absent the genovar score will be 1,4,[5],12:i:-.

4. Other overlaps include:

   a) S. Blegdam with S. Enteritidis (known overlap)
   b) S. Nitra with S. Enteritidis (known overlap)
   c) S. Kiel with S. Dublin (known overlap)
   d) S. Rostock with S. Dublin (known overlap)
   e) S. Kimuneza with S. Bredeney (potential overlap – seen within the AHVLA data set)
   f) S. Nagoya with S. Telelkibir (potential overlap – seen within the AHVLA data set)
   g) S. Sundsrall with S. Telelkibir (potential overlap – seen within the AHVLA data set)

   Potential overlaps are those seen in the AHVLA data set, it may be possible to adjust the software to correct theses as in some instances the software did not look at the secondary probe set.

- A detailed knowledge of *Salmonella* serotypes would still play a role in deciding where the result obtained could be a correct result even when another serotype (or description) is proposed by C&T. Although the routine testing could be performed by personnel with little knowledge of *Salmonella*.

Partially correct structures proposed – 2.06%

- 2.06 % of C&T results were partially correct but would be inaccurate for a diagnosis, this figure includes strains that were partially typed using the classic KW scheme, where C&T gave the nearest match, e.g. O Rough:G,M: was typed as S. Enteritidis by C&T. It is possible that the gene may be present in this isolate but not expressed and therefore not detected by classical serotyping methods. PCR and sequencing of amplicon would be required to verify the presence of the gene in this isolate.

- These false positive results are less of a concern as before any recommendation to farmers or culling of flocks a full KW confirmatory assay would be performed.

Positive Salmonella not detected – 0%

- 0% of the tests tested positive for no *Salmonella*. All *Salmonella* samples tested positive for *Salmonella*, no negative samples obtained. 6 *Salmonella* strains gave a No *Salmonella* result but these were all due to reaction failures where other control spots were not identified, therefore cannot be included. In addition a panel of 24 non-*Salmonella* isolates were included as negative *Salmonella* controls and included Escherichia coli, Staphylococcus epidermidis and Staphylococcus aureus, all gave a No *Salmonella* result.

Test failure repeat

- The reproducibility of C&T was assessed by repeat analysis of isolates. Where there was disparity with KW serotyping the sample was retested with C&T on different days. In addition, a selection of matching results was also retested to ensure repeatability. A total of 99 isolates were repeated for various reasons, including reaction failures and human errors. Where possible errors were retested using the same DNA preparation then if the sample failed again then a fresh DNA preparation from the original culture collection was obtained. The isolates that did not match the proposed serotype after repeating with a fresh DNA stock then it is strongly advised that these strains are re-serotyped using the classic KW scheme as the mismatches are between samples that are “sero-typically” different. Without repeating the KW serotyping on the isolate it is not possible to say whether there is an issue with the C&T method or an error has been made with the classic KW scheme.

Automated detection

- Automated detection. The transfer to automation was performed using 228 isolates. Of these 228 automated reads, a total of 31 were also assessed by manual detection – all 31 automated and manual genovar scores were a 100% match.

Conclusions

The SGSA and the C&T array have both demonstrated their ability to provide a rapid molecular based detection of *Salmonella* serotypes; both have the potential to easily transfer to an automated platform to allow for minimal human hands-on time / error in the array; neither array hit the 99.9% for both sensitivity and specificity for *S.* Enteritidis and *S.* Typhimurium as determined by the EU. Neither array produced a false negative for these serotypes only false positive identification and it would be recommended that if either array was used as a rapid alerting system then a temporary action could be placed on a farm with no further action on farmers without a confirmatory KW assay. It should be pointed out that all *Salmonella* isolates were tested only once by the classic KW serotyping scheme and it has been assumed rather than confirmed that all serotypes called by classic serotyping were correct. This is not an unreasonable position from which to assess array platforms. However, there is evidence from the testing that perhaps some KW serotyping required re-checking.

The C&T array is available commercially in kit form (including the DNA extraction) and has an established analysis software package with a constantly updated database and pro-active support team in The Netherlands. Thus, little knowledge of *Salmonella* is required to interpret the data produced. The design of the C&T array means that even *Salmonella* serotypes not currently recognised by the software can be determined

as *Salmonella*, assigned a Genovar and submitted to Checkpoints for inclusion into the database and software. These are under constant review by the Checkpoints team. In theory the SGSA could perform to the same level of detection for up to 1358 of the known serotypes but has only been validated for the 59 stated.

The C&T array has more "evidence" for the results obtained due to the larger sample number included but it must be noted that unless 4603 isolates of each S. Enteritidis and S. Typhimurium were detected by either array without an error then the results cannot be considered truly statistically significant due to the high EU regulation requirements of detection for these two serotypes; ≥99% at 99% confidence levels.

The developmental stage of the C&T array makes it the advantageous choice for early selection as a rapid alerting system. The array is already approved by the AOAC-RI and OIE for specified serotypes (see table 9) and the array has had a robust validation at the AHVLA on 171 different serotypes.

Table 12 – Broad overview of the SGSA and C&T array

Comparison of SGSA and C&T array		
	SGSA	C&T
NO. of Serotypes detected	59	300+
Accreditation(s)	No	Yes
No. of samples automation per run	96	36
Detection of SE & ST	99.7%	100%
Detection of all serotypes stated	95.2%	97.84%
Commercially available	No	Yes

This study has been very robust especially in the use of the defined panel of 104 serotypes. This collection comprises core serotypes but also extremely rare serotypes, such as S. Nitra that have never been reported in the UK or Europe for over four decades, and serotypes such as S. Rostock and S. Begdam that are not only rare but belonging to group D salmonellas are known to be very difficult to type by the classic KW serotyping method. Given the apparent success of the C&T system for detection of core serotypes *S. Typhimurium* and *S. Enteritidis* (100% success rate) it can be argued that arrays have the potential in certain key facets of Salmonella diagnosis play a leading role. An output from this study is validation that arrays do perform well and that there is a need to perhaps consider the needs of serotyping in the UK legislative context. It is possible that cost benefit analysis of the application of arrays supported by KW serotyping may yield improved reporting times and significant cost savings. This is potential future work.

References to published material

9. This section should be used to record links (hypertext links where possible) or references to other published material generated by, or relating to this project.

1. **Anonymous (2010a)** The Community Summary Report on Trends and Sources of Zoonoses, Zoonotic Agents and Food-borne Outbreaks in the European Union in 2008, EFSA Journal; 2010 8(1):1496.
2. **Anonymous (2010b)** EFSA Panel on Biological Hazards (BIOHAZ); Scientific Opinion on monitoring and assessment of the public health risk of "*Salmonella* Typhimurium-like" strains. EFSA Journal 8(10):1826.
3. **Franklin *et.al.* (2011)** Rapid genoserotyping tool for classification of *Salmonella* serovars. J Clin Microbiol. 49(8):2954-65.
4. **Grimont and Weill (ed.). 2007**. Antigenic formulae of the *Salmonella* serovars, 9th ed. WHO Collaborating Centre for Reference and Research on *Salmonella*. Institut Pasteur, Paris, France. http://www.pasteur.fr/ip/portal/action/WebdriveActionEvent/oid/01s-000036-089.
5. **Galanis *et.al.* (2006)** Web-based surveillance and global *Salmonella* distribution, 2000-2002. Emerg Infect Dis. 12(3):381-8.
6. **Hendriksen *et.al.* (2011)** Global monitoring of *Salmonella* serovar distribution from the World Health Organization Global Foodborne Infections Network Country Data Bank: results of quality assured laboratories from 2001 to 2007. Foodborne Pathog Dis. 8(8):887-900.
7. **Herikstad *et.al.* (2002)** *Salmonella* surveillance: a global survey of public health serotyping. Epidemiol Infect. 129(1):1-8.

8. **Franklin K**, Lingohr EJ, Yoshida C, Anjum M, Bodrossy L, Clark CG, Kropinski AM, Karmali MA. Rapid Genoserotyping Tool for Classification of *Salmonella* Serovars J Clin Microbiol. 2011 Aug;49(8):2954-65.