

RESEARCH ARTICLE

Open Access



# Whole genome sequencing of *Salmonella* Typhimurium illuminates distinct outbreaks caused by an endemic multi-locus variable number tandem repeat analysis type in Australia, 2014

Anastasia Phillips<sup>1\*</sup>, Cristina Sotomayor<sup>1,2</sup>, Qinning Wang<sup>1,2,3</sup>, Nadine Holmes<sup>1,2</sup>, Catriona Furlong<sup>4</sup>, Kate Ward<sup>4</sup>, Peter Howard<sup>3</sup>, Sophie Octavia<sup>5</sup>, Ruiting Lan<sup>5</sup> and Vitali Sintchenko<sup>1,2,3</sup>

## Abstract

**Background:** *Salmonella* Typhimurium (STM) is an important cause of foodborne outbreaks worldwide. Subtyping of STM remains critical to outbreak investigation, yet current techniques (e.g. multilocus variable number tandem repeat analysis, MLVA) may provide insufficient discrimination. Whole genome sequencing (WGS) offers potentially greater discriminatory power to support infectious disease surveillance.

**Methods:** We performed WGS on 62 STM isolates of a single, endemic MLVA type associated with two epidemiologically independent, food-borne outbreaks along with sporadic cases in New South Wales, Australia, during 2014. Genomes of case and environmental isolates were sequenced using HiSeq (Illumina) and the genetic distance between them was assessed by single nucleotide polymorphism (SNP) analysis. SNP analysis was compared to the epidemiological context.

**Results:** The WGS analysis supported epidemiological evidence and genomes of within-outbreak isolates were nearly identical. Sporadic cases differed from outbreak cases by a small number of SNPs, although their close relationship to outbreak cases may represent an unidentified common food source that may warrant further public health follow up. Previously unrecognised mini-clusters were detected.

**Conclusions:** WGS of STM can discriminate foodborne community outbreaks within a single endemic MLVA clone. Our findings support the translation of WGS into public health laboratory surveillance of salmonellosis.

**Keywords:** *Salmonella* Typhimurium, Molecular epidemiology, Foodborne outbreaks, Whole genome sequencing, Public health surveillance

## Background

*Salmonella* gastroenteritis is responsible for considerable disease burden in both developed and developing countries, with an estimated 93.8 million cases and 155,000 deaths each year [1]. The majority of these cases are foodborne [1–3], although transmission may occur directly from infected persons [4]. *Salmonella* is frequently identified as the aetiological agent in foodborne

outbreaks in Australia, with *Salmonella enterica* subsp. *enterica* serovar Typhimurium (*S.* Typhimurium/STM) being the predominant serovar [5].

Timely subtyping of *Salmonella* and integration of laboratory findings into public health actions are critical in reducing delays in outbreak investigation [6]. In New South Wales (NSW), Australia, where STM accounts for around half of all culture-confirmed cases of salmonellosis [7], multi-locus variable number tandem repeat analysis (MLVA) has been used since 2006 to prospectively subtype STM and identify potential clusters. MLVA measures the variable length of five STM loci and has

\* Correspondence: Anastasia.Phillips@health.nsw.gov.au

<sup>1</sup>Centre for Infectious Diseases and Microbiology-Public Health, Westmead Hospital, Sydney, NSW, Australia

Full list of author information is available at the end of the article

been considered faster and more discriminatory than other historical typing methods [7, 8]. However, common or endemic MLVA types may cause multiple outbreaks along with sporadic cases and a few selected MLVA types may represent a large portion of isolates observed in a geographical location [9].

Whole genome sequencing (WGS) has offered the ultimate discriminatory power with the potential to enhance epidemiological investigations and elucidate transmission pathways [9–11]. As WGS continually becomes quicker and more cost effective, and data quality improves [12–15] its potential to be used in routine epidemiological typing increases [15]. WGS has been recently used to understand outbreak sources and transmission patterns of other diseases, including enteric diseases (such as *Shigella sonnei* and enterohaemorrhagic *Escherichia coli*) [12, 14, 16–18]. Furthermore, WGS has the potential to discriminate between sporadic and outbreak isolates which may be indistinguishable by current methods of subtyping [14, 17, 19, 20].

Recently, the potential of WGS for characterising *Salmonella* outbreaks and differentiating outbreak and sporadic strains has been explored [15, 21] and the importance of correlating WGS and epidemiological data has been emphasised [15]. With increased understanding of the significance of genomic variability, there is potential to progress to near real-time, discriminatory genotyping of outbreak cases [21].

The primary aim of this study was to examine the utility of WGS in discriminating sporadic and outbreak-linked STM infections within the same endemic MLVA type. This has the potential to improve the accuracy of cluster definitions and timeliness of outbreak response. We report the WGS analysis of human and food STM isolates from two epidemiologically independent outbreaks of a single MLVA type in NSW during 2014, along with interspersed sporadic cases.

## Methods

### Isolate selection

Salmonellosis is a notifiable disease in Australia. Details of cases are maintained in the NSW jurisdictional Notifiable Conditions Information Management System (NCIMS) Database by the Communicable Disease Branch, Health Protection NSW. All case isolates presumptively identified as *Salmonella* by pathology service providers in NSW are forwarded to the NSW State Enteric Reference Laboratory at the Centre for Infectious Diseases and Microbiology, Institute for Clinical Pathology and Medical Research-Pathology West (ICPMR) in Sydney. Isolates undergo confirmatory testing, serotyping and MLVA-5 typing at ICPMR. Relevant food isolates from outbreak investigations are also typed and stored.

A common MLVA type (2-15-9-10-0212 or 3-17-10-11-523 according to the European CDC [22] and Australian conventions [23], respectively) that was known to have caused two food-borne outbreaks along with sporadic cases between January and May 2014 in metropolitan Sydney was selected for this study. STM of this MLVA type was responsible for 3.7 % of all STM infections in 2014. These isolates belonged to phage type DT135. The two outbreaks were dispersed in time and space such that sporadic cases occurred within two months of either or both outbreaks.

Each case isolate used in this study was classified as “outbreak”, “sporadic” or “secondary” by an epidemiologist at the Communicable Disease Branch based on information that was available for each case in NCIMS. This information was a case interview, a completed online questionnaire, and/or demographic information. Outbreak cases were those with documented exposure at one of the established outbreak venues. Sporadic cases were those with no documented exposure to one of these venues in the time frame that fitted with the incubation period for *Salmonella*. Secondary cases were defined as those with documented household contact with a confirmed case of the same MLVA type within a time frame that fitted with the incubation period for *Salmonella*. Environmental isolates cultured from food samples or surface swabs during the public health investigation of the two outbreak venues were also included in the analysis. Background information on the two outbreaks was provided by the Communicable Disease Branch and the NSW Food Authority. Isolates were labelled according to a study number and partial laboratory accession number.

### DNA sample preparation

A single colony of each STM isolate was sub-cultured on blood agar plates at 37 °C for 24 h. DNA extraction for MLVA-5 was prepared using a boiling method [23]. For whole genome sequencing, the DNA was extracted and purified using a DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer’s instructions. DNA quantities were estimated using the Qubit dsDNA HS Assay Kit and the Qubit Fluorometer (Life Technologies, Germany) according to the manufacturer’s instructions.

### MLVA-5

All isolates were MLVA typed using 5 VNTR loci (referred to as MLVA-5) based on an original method by Lindstedt et al. [24] with modifications as described in Wang et al. [23], compatible to the European CDC Scheme [22].

### Genome sequencing

A 100 bp paired-end library (200 cycles) was prepared for each purified DNA sample using the NexteraXT kit (Illumina). Each of the DNA samples was barcoded

using the Nextera XT index kit v2 Set A (Illumina) then pooled and sequenced on the HiSeq platform (Illumina, Inc, San Diego, CA, USA) using a single lane of a HiSeq flow cell. Quality control of the libraries was performed by assessing library size distribution on an Agilent 2200 TapeStation (fragment sizes typically range between 250–1000 bp) and libraries were quantified by real time PCR using the KAPA library quantification kit according to the manufacturer's protocol (Kapa Biosystems). Raw short read sequence data was submitted to Sequence Read Archive (SRA) at the National Centre for Biotechnology Information (NCBI) under the SRA accession number SRP074336.

### SNP analysis

FastQ files were imported into CLC Genomics Workbench v 7.0 (CLC bio, Aarhus, Denmark) and reads were trimmed to remove Nextera transposase adapter sequences then mapped to the reference genome of *S. Typhimurium* LT2 (NCBI GenBank Accession No. NC\_003197). Quality-based variant detection was performed using settings of a minimum neighborhood quality of 15 and minimum central quality of 20. Variant detection thresholds were set for a minimum coverage of 10 reads and minimum variant frequency of 75 %. *De novo* assembly was performed on reads from one of the study isolates (isolate 65\_E2330) using CLC Genomics Workbench and the resulting contigs were scaffolded using the CLC Microbial Genome Finishing Module. The consensus sequence was extracted from the scaffolded contigs and was then used as a second reference genome for read mapping and variant detection, as described above, to validate SNPs. All predicted SNPs were visually examined by locating the SNP positions in the read-mapping files and comparing the consensus calls to the reference genome. Gene and amino acid allele changes associated with each SNP were determined from the annotated reference genome. We also compared the SNP calls with the method used by us previously [25] and found that the two methods gave highly concordant results with only 2 SNPs not detected in our assembly based method. However, CLC Workbench offers potentially better reproducibility.

### Ethics approval

This study was approved by the Western Sydney Local Health District Human Research Ethics Committee (LNR/14/WMEAD/235, granted 21 July 2014).

## Results

### Isolate selection and outbreak investigation

Of 85 isolates of MLVA 2-15-9-10-0212 recovered from human cases diagnosed between January and May 2014 (inclusive), 56 isolates were epidemiologically classified

from the available data and were sequenced in this study. One case was classified as secondary to an outbreak case. A further 10 environmental isolates from the two outbreaks were also included for sequencing. One environmental isolate (65\_E2330) from outbreak M had a slightly different but related MLVA profile 2-15-9-9-0212. An isolate with one repeat difference can be regarded as part of this cluster due to rapid evolution at this locus [26].

Following whole genome sequencing, two case isolates from Outbreak A and two from Outbreak M were excluded from the analysis due to the poor quality of sequences. The total number of case isolates included in the study was 52.

Outbreak A was linked to chicken liver pate produced weekly from fresh ingredients at a Sydney café. The contaminated ingredient was likely to have been brought in on or around the 25<sup>th</sup> of January 2014 as a single event [Food Authority NSW, personal communication, 14 August 2014]. Outbreak M appeared to be caused by contamination of a number of foods and environmental surfaces at a Sydney hot bread shop. All sporadic cases occurred within two months of one or both outbreaks (Fig. 1).

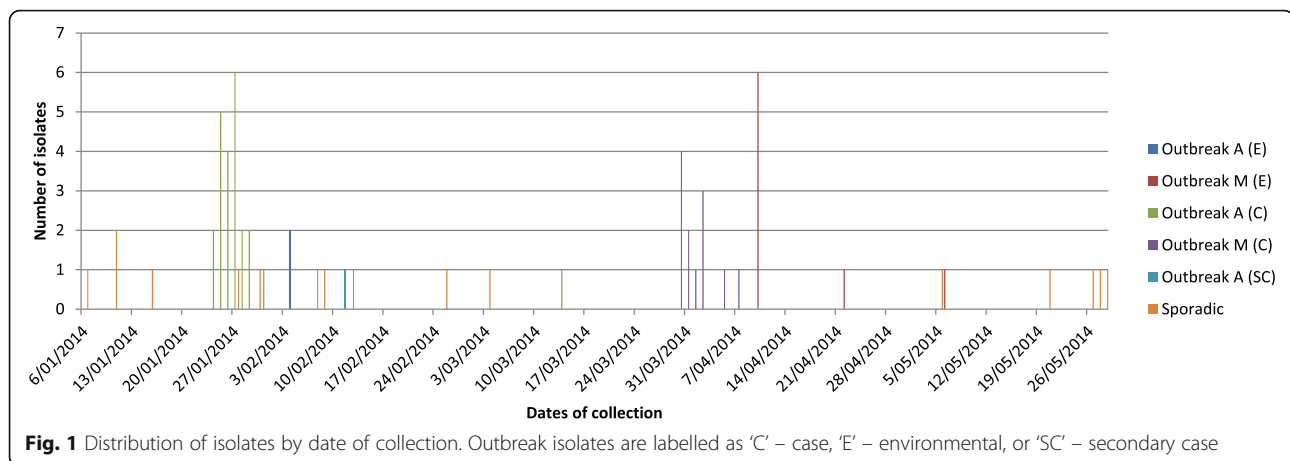
### Whole genome sequencing

#### SNP identification

The genomes were sequenced in one multiplex with an average depth of coverage of 80X ranging from 37X to 154X. A total of 30 SNPs were identified from the 62 isolates. The SNPs called were identical independently of whether the reads were mapped to STM LT2 or to the *de novo* assembly of isolate 65. Based on SNP profiles, all isolates were clustered into groups consistent with their known epidemiological sources (Fig. 2). All case isolates from outbreak M had a unique non-synonymous SNP (C to T) at position 723663 (reference STM LT2 base position), which was also present in the environmental and food isolates obtained from outbreak M. All of the isolates from outbreak M also had a single nucleotide insertion (A) at position 1789781, apart from one case isolate (isolate 44\_M4945). In addition, two isolates from outbreak M (one case isolate 42\_M5132 and one environmental isolate 61\_E4721) carried unique genomic variations at positions 2162287 (a deletion) and 3983630 (a synonymous SNP), respectively.

All of the case and environmental isolates from outbreak A shared a unique single nucleotide insertion (A) at position 7408. One of the case isolates (isolate 11\_A2137) also had a unique SNP at position 185058.

The effect of genomic changes on translation was variable. The genomic features of SNPs identified, including the SNP position and type, allele length, annotation,



coding region and amino acid changes are provided in the supplementary Additional file 1: Table S1.

Along with SNP differences described above, a seven base-pair deletion at position 2332558 was identified (Fig. 2). This deletion was present in all of isolates from Outbreak M but not those from Outbreak A. The deletion is in *oafA* gene which affects the function of the somatic O antigen of STM – those with the deletion were O antigen factor v (1, 4, 5, i, 1, 2) and those without were O antigen factor xii (1, 4, 12, i, 1, 2) [27].

Sporadic case isolates showed distinct SNP profiles compared to the isolates from Outbreak A and M (Fig. 2). Around half had two additional SNPs (A to C at position 3491775 and 4273259) and there were several unique SNPs randomly distributed among the sporadic isolates. Four cases were distinguished by only one SNP from Outbreak A (isolates 1\_N4472, 22\_N2054, 31\_N4583 and 32\_N5047). On examining the epidemiological data, these cases were not linked to Outbreak A. Case 1\_N4472 had an onset date well before the introduction of a contaminated ingredient to the Outbreak A venue and 22\_N2054 confirmed no link with the venue on interview. Neither 31\_N4583 nor 32\_N5047 stated any link to Outbreak A during online survey. One of the sporadic isolates (54\_N4956) had the same insertion at position 1789781 as the Outbreak M isolates and also had the seven base-pair deletion. However, this isolate also had a SNP at position 3413980 and did not have the other SNPs characteristic of Outbreak M (723663).

Among the sporadic isolates, there were mini-clusters detected by WGS analysis (Fig. 2). One of these clusters contained four cases (3\_N2463, 29\_N4955, 24\_N4802 and 36\_N5175), all of which had a common SNP and shared the seven base-pair deletion. The cases occurred over a period of 52 days between January and March 2014. The later three cases shared an additional common SNP and all resided in one geographical area in metropolitan Sydney. Cases in another cluster (52\_N5064 and 53\_N5139a) shared

three common SNPs and were clustered in time (occurring six days apart). In the last cluster (37\_N4954 and 67\_N4799), the isolates were indistinguishable by SNP analysis, but appeared genetically variable because of the presence of the seven base-pair deletion (different strains with different O antigens). These cases were also separated in time by seven weeks and did not reside in the same geographical area.

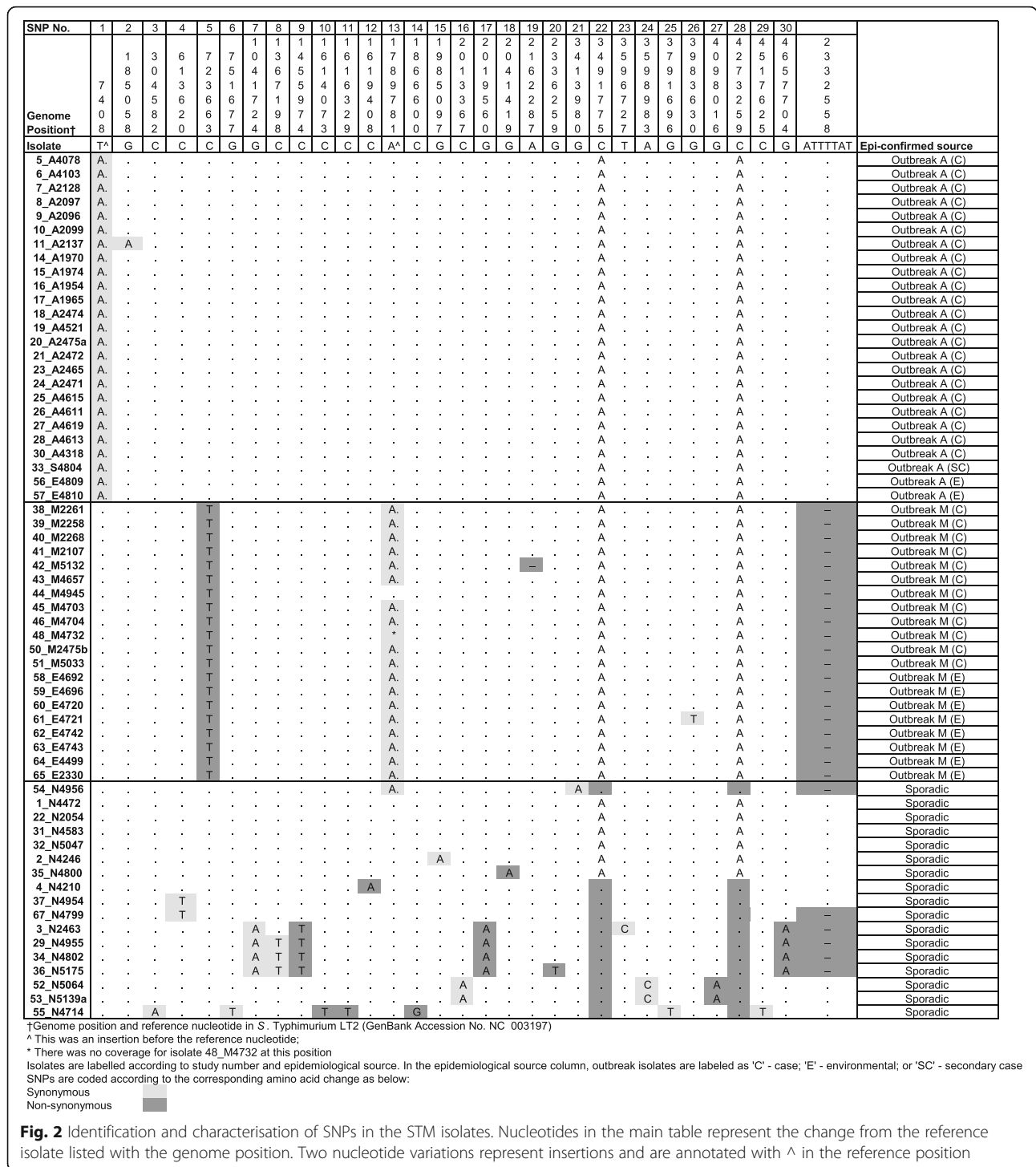
Concatenated SNP profiles were further analysed using a minimum spanning tree (MST) approach (Fig. 3). The MST demonstrated that the two outbreaks were clustered separately and that most sporadic cases were more than two SNPs away from the outbreaks and from each other, with only a few exceptions. The four isolates cluster separated by one SNP from outbreak A is actually two nucleotide polymorphism variations when the 7-base pair deletion is included (Fig. 2).

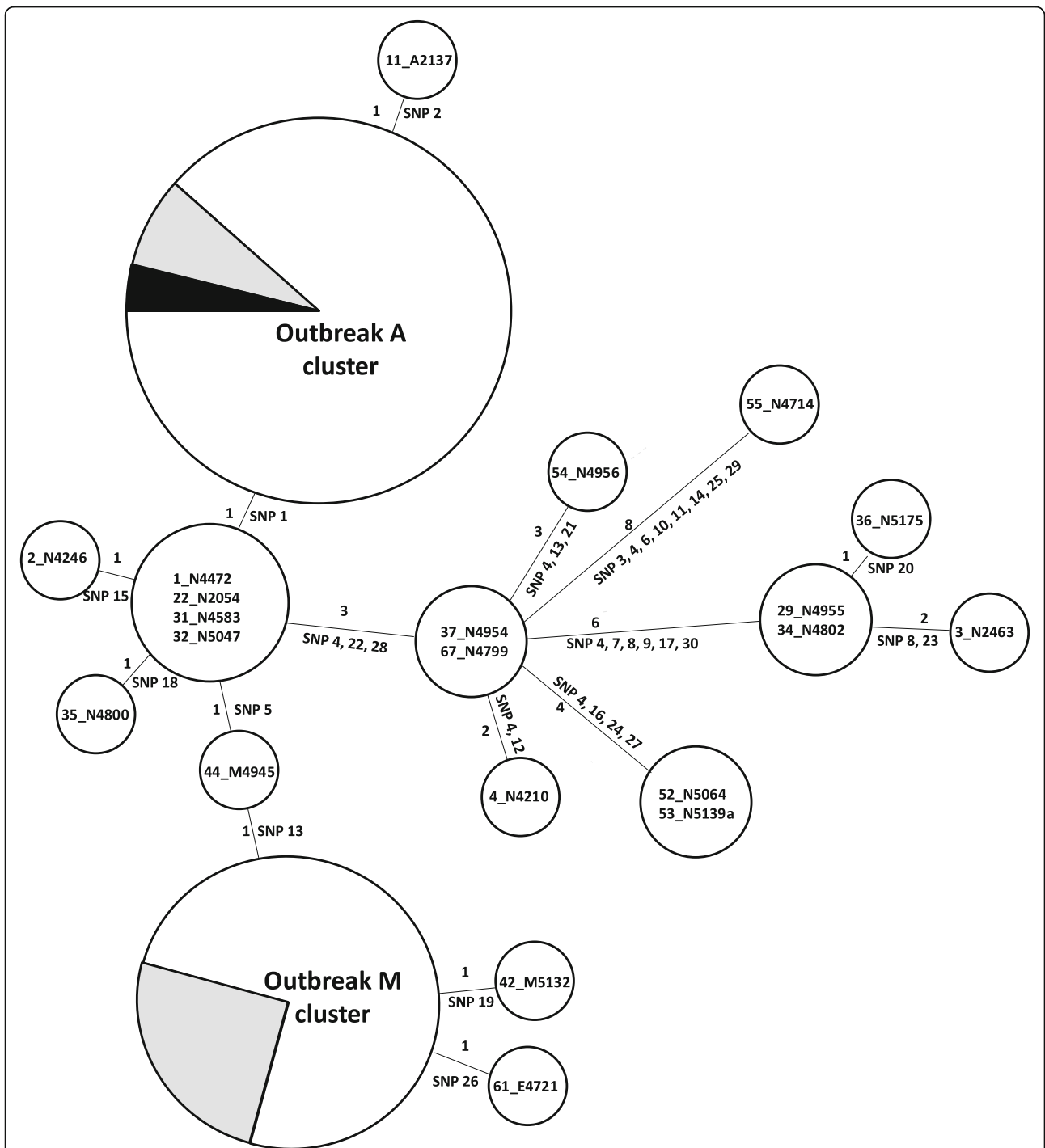
The phylogeny of the outbreak strains was considered in the context of national and international STM isolates. We included 137 isolates from other studies including 75 Australian isolates and 62 global isolates for comparison [15, 21, 25, 28–30]. The phylogenetic tree (Additional file 2: Figure S1) showed that the isolates from this study were tightly clustered together as one group.

## Discussion

Our study examined a collection of STM isolates representing two epidemiologically distinct outbreaks in an urban area along with spatiotemporally associated sporadic cases that were initially linked by MLVA typing. WGS clearly differentiated these isolates in concordance with epidemiological evidence, providing greater discrimination than MLVA-5 alone.

Our study demonstrated unique genomic variations in each of the outbreak clusters, with one or two specific nucleotide variations in each outbreak group. This is





**Fig. 3** Minimum spanning tree of SNPs identified from the 62 STM isolates. Each circle represents isolates with indistinguishable genome (by SNPs excluding indels) and the size of circles corresponds to the number of isolates. Numbers above or on the left of the connecting lines between circles are the number of SNPs while the SNPs as numbered in Fig. 2 were listed either below or on the right of the connecting lines. Labelling of isolates is consistent with Fig. 2. The grey colour in the outbreak circles indicates the proportion of environmental isolates and the black colour indicates the isolate from the secondary case. Outbreak A cluster included 24 out of 25 isolates (except 11) and Outbreak M cluster included 17 of 20 of the outbreak isolates (except 42, 44, and 61) shown in Fig. 2

Our isolates were all of identical MLVA type and recovered from residents within a confined geographic area of metropolitan Sydney, demonstrating the superior discrimination offered by WGS over current MLVA-5 typing in highly clonal isolates. However, our previous study showed that isolates from the same outbreak may differ by up to 4 SNPs [25]. The low number of genetic differences between outbreak isolates in this study suggests that the two outbreaks may share a common source of origin, such as a single egg manufacturer. The number of SNP difference between isolates observed depends on mutation rate and the evolutionary time that has passed. The mutation rate estimates for *S. Typhimurium* are varied. The lowest rate is  $1.9 \times 10^7$  substitutions per site per year estimated from ST313 causing invasive infections in Africa, the intermediate rate is  $3.4 \times 10^7$  substitutions per site per year from epidemic DT104 infections, and the highest rate is  $12 \times 10^7$  substitutions per site per year from a DT135a outbreak [21, 31, 32]. These rates correspond to an accumulation of approximately one to five SNPs per genome per year. Therefore, the small number of SNP differences observed between our isolates suggests that it is highly likely that these isolates shared a very recent common ancestor.

Other studies have found highly similar isolates within individual outbreaks, along with greater variability between outbreaks, where eggs from the same source were implicated [21, 33]. In NSW, investigation of eggs farms is conducted by the NSW Food Authority and includes MLVA typing of identified *Salmonella* strains. In future investigations, WGS is likely to offer further insights that could make public health investigations more targeted and may allow trace back investigations for a possible common source for outbreaks with closely related isolates such as the outbreaks examined in this study.

Our case outbreak isolates were also largely indistinguishable from the environmental isolates from the same outbreak. Many of our sporadic cases were also well differentiated from the outbreak cases by WGS, and did not carry the unique outbreak variations. However, some sporadic cases showed only minimal differences from the outbreak cases, including the four cases that differed by only one SNP from outbreak A. In addition, one sporadic isolate had the outbreak M insertion and several sporadic isolates contained the 7-bp deletion.

The difference between the outbreak and sporadic isolates was less pronounced than in the previous studies, which demonstrated difference of between 75–100 SNPs [21] to up to several thousand [34] SNPs between sporadic and outbreak isolates. While previous studies included isolates with greater underlying variability, including STM of differing MLVA types collected over a longer time period [21, 34], the small

SNP differences between some of our sporadic and outbreak isolates again suggest the possibility of a common environmental (i.e. farm) source. While our sporadic cases clustered separately from outbreak isolates by WGS, the small differences highlight the importance of interpreting epidemiological data alongside WGS findings. The WGS findings suggest that further public health investigation and trace back may be necessary to identify the potential source of both outbreak and sporadic cases.

Unexpectedly, within our sporadic cases, WGS grouped isolates into several apparent and previously undetected mini-clusters. For one of these clusters, epidemiological data indicated the cases were geographically linked. For the others, while the sequencing of these genomes suggested that they were related to each other, public health investigation have not been able to identify any common links between the cases. These findings highlight the need for further comparisons of clusters defined by genomic and public health surveillance and the role of sporadic cases in the natural evolution of outbreaks.

Our report adds to the emerging evidence that WGS challenges traditional epidemiological and typing approaches in the investigation of foodborne outbreaks. For example, previous studies of *Salmonella* Enteritidis and STM suggested the ability of WGS to detect previously unidentified clusters [35] and in a recent study of *Shigella sonnei*, WGS differentiated apparent outbreak strains into three genomically distinct clusters and revealed misclassification by traditional methods [17]. In another recent study in NSW, WGS was able identify STM isolates that were not originally considered to be part of the outbreak [25].

However, genomic surveillance provides only one line of evidence [15] which has to be synthesised with epidemiologic approaches to identify suspected sources [21]. This method, described as 'the ultimate stain typing system' [13], has the potential to improve the definition of community outbreaks and assist in a more rapid and targeted public health response through accurate microbiological discrimination of outbreak cases. Recently, a study in the US suggested that WGS of *Salmonella* isolates was feasible for prospective surveillance [35]. In outbreaks of gastroenteritis caused by enterohaemorrhagic *Escherichia coli* (EHEC) in Europe, WGS enabled tracking of outbreak isolates in real-time through the identification of four informative SNPs unique for outbreak isolates [18] and through web-based bioinformatics analysis [14]. The United States Food and Drug Administration, in collaboration with the National Centre for Biotechnology Information and public health laboratories are trialling a centralised surveillance system for WGS data [36]. However, significant challenges in

bioinformatics infrastructure, data governance and training of epidemiologists have to be addressed to fully realise benefits of genomics-guided improvements in the delivery of public health surveillance and response [37–39]. A robust and standardised approach to identify potential outbreaks needs to be developed based on the range of SNP differences and the likelihood of those resulting from a recent common ancestor.

A limitation of our study was that we relied on retrospectively well-defined outbreaks. Our epidemiological data, while robust, did not provide sufficient detail for some sporadic cases, which emerged as potentially clustered only following WGS.

## Conclusions

Whole genome sequencing of STM associated with acute gastroenteritis can illuminate distinct foodborne community outbreaks amongst groups of isolates sharing the same MLVA profile. Genomic analyses can also further differentiate sporadic from outbreak cases and cluster sporadic isolates within endemic MLVA types of STM, which can significantly improve the resolution of public health laboratory surveillance. These findings strengthen the case for the adoption of WGS-guided public health surveillance of human salmonellosis and outlines the challenges of distinguishing sporadic cases with minimal differences from outbreak cases.

## Additional files

**Additional file 1: Table S1.** Genomic features of SNPs identified. (XLS 33 kb)

**Additional file 2: Figure S1.** Phylogeny of the outbreak A and M strains in the context of national and international STM isolates. Genome data analysed in Octavia et al. representing five STM outbreaks in Australia [25]; Kingsley et al. representing ST313 outbreak in Malawi [30]; Leekitcharoenphon et al. representing six STM outbreaks in Denmark [15] and Hawkey et al. representing STM DT135a outbreak in Australia [21] were also included as comparisons and marked as the corresponding study/outbreak. Other branches that are not labelled are background isolates from the above studies; draft genomes from Pang et al. [29] which include five diverse Australian STM isolates; Fu et al. representing *Salmonella* reference collection A; [28] and other fully sequenced STM genomes available from GenBank including LT2 (Accession No. NC003197), 798 (Accession No. CP003386), DT2 (Accession No. HG326213), DT104 (Accession No. HF937208), 14028S (Accession No. CP001363), SL1344 (Accession No. FQ312003), UK-1 (Accession No. CP002614), T000240 (Accession No. AP011957), U288 (Accession No. CP003836) and ST4/74 (Accession No. CP002487). Bootstrap values if greater than 50 %, are presented on the internal branches. (PPTX 74 kb)

## Abbreviations

ICPMR: Institute for Clinical Pathology and Medical Research; NCIMS: Notifiable Conditions Information Management Systems; SNP: Single nucleotide polymorphism; STM: *Salmonella* Typhimurium; WGS: Whole genome sequencing

## Acknowledgements

The authors thank Dr Joey Lai from Westmead Millennium Institute Genomics Facility for performing genomic library preparation using

NexteraXT kit and the Microbiological Diagnostic Unit, The University of Melbourne for phage typing of STM isolates; and Mr Craig Shadbolt, NSW Food Authority, for background information on the outbreaks. This study was partly supported by a grant from the National Health and Medical Research Council of Australia (APP1050227).

## Funding

This project was supported by the Centre for Infectious Diseases and Microbiology (Public Health) and grants from the National Health and Medical Research Council of Australia (APP1044986 and APP1050227). The funding body had no influence over the design of the study or collection, analysis, or interpretation of data or in writing the manuscript.

## Availability of data and materials

The dataset analysed during the current study is available in the Sequence Read Archive (SRA) (GenBank, National Centre for Biotechnology Information) under the repository, <http://www.ncbi.nlm.nih.gov/sra/?term=SRP074336> [SRA accession: SRP074336].

## Authors' contributions

AP designed the protocol, sought ethics approval, selected cases, conducted descriptive analysis, conducted SNP analysis and drafted the manuscript. CS identified study isolates and conducted DNA extraction. QW supervised DNA extraction and WGS, conducted SNP analysis and assisted in drafting manuscript. NH conducted SNP analysis and assisted in drafting manuscript. CF provided epidemiological data and applied classifications and assisted with manuscript. KW assisted with ethics application and data access. PH provided confirmatory testing, serotyping and MLVA typing of isolates. SO and RL conducted further data analysis and assisted with drafting the manuscript. VS supervised protocol design and analysis of findings, and assisted in drafting manuscript. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

Not applicable.

## Author details

<sup>1</sup>Centre for Infectious Diseases and Microbiology–Public Health, Westmead Hospital, Sydney, NSW, Australia. <sup>2</sup>Marie Bashir Institute for Emerging Infectious Diseases and Biosecurity and Sydney Medical School, The University of Sydney, Sydney, NSW, Australia. <sup>3</sup>NSW Enteric Reference Laboratory, Centre for Infectious Diseases and Microbiology Laboratory Services, Pathology West, Sydney, NSW, Australia. <sup>4</sup>OzFood Net, Communicable Disease Branch, Health Protection, NSW Ministry of Health, Sydney, NSW, Australia. <sup>5</sup>School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW, Australia.

Received: 23 January 2016 Accepted: 6 September 2016

Published online: 15 September 2016

## References

- Majowicz S, Musto J, Scallan E, Angulo F, Kirk M, O'Brien S, et al. The Global Burden of Nontyphoidal *Salmonella* Gastroenteritis. *Clin Infect Dis*. 2010;50(6):882–9.
- Australian Institute of Health and Welfare. Australia's health 2012. Canberra (AUST): AIHW2012 Contract No.: Cat. no. AUS 156.
- Greig J, Ravel A. Analysis of foodborne outbreak data reported internationally for source attribution. *Int J Food Microbiol*. 2009;130:77–87.
- King N, Lake R, Campbell D. Source Attribution of Nontyphoid Salmonellosis in New Zealand Using Outbreak Surveillance Data. *J Food Prot*. 2011;74(3):438–45.
- The OzFoodNet Working Group. OzFoodNet enhanced foodborne disease surveillance, 1 January to 31 March 2013. *CDI*. 2014;38(1):E70-77.



6. Hedberg C, Greenblatt J, Matyas B, Lemmings J, Sharp D, Skibicki R, et al. Timeliness of Enteric Disease Surveillance in 6 US States. *Emerg Infect Dis*. 2008;14(2):311–3.
7. Wang Q, Chiew R, Howard P, Gilbert G. *Salmonella* typing in New South Wales: current methods and application of improved epidemiological tools. *NSW Public Health Bull*. 2008;19(1–2):24–8.
8. Gilbert G. Using MLVA to type strains of *Salmonella* Typhimurium in New South Wales. *NSW Public Health Bull*. 2008;19(1–2):29–31.
9. Sintchenko V, Wang Q, Howard P, Ha C, Kardamanidis K, Musto J et al. Improving resolution of public health surveillance for human *Salmonella enterica* serovar Typhimurium infection: 3 years of prospective multiple-locus variable-number tandem-repeat analysis (MLVA). *BMC Infect Dis*. 2012;12(78). doi:10.1186/1471-2334-12-78.
10. Köser C, Holden M, Ellington M, Cartwright E, Brown N, Ogilvy-Stuart A, et al. Rapid Whole-Genome Sequencing for Investigation of a Neonatal MRSA Outbreak. *NEJM*. 2012;366:2267–75.
11. Le V, Diep B. Selected insights from application of whole-genome sequencing for outbreak investigations. *Curr Opin Crit Care*. 2013;19(5):432–9.
12. Didelot X, Bowden R, Wilson D, Peto T, Crook D. Transforming clinical microbiology with bacterial genome sequencing. *Nat Rev Genet*. 2012;13(9):601–12.
13. Gilbert G, Sintchenko V. The use of mycobacterial interspersed repetitive unit typing and whole genome sequencing to inform tuberculosis prevention and control activities. *NSW Public Health Bull*. 2013;24(11):10–4.
14. Joensen K, Scheutz F, Lund O, Hasman H, Kaas R, Nielsen E, et al. Real-Time Whole-Genome Sequencing for Routine Typing, Surveillance, and Outbreak Detection of Verotoxigenic *Escherichia coli*. *J Clin Microbiol*. 2014;52(5):1501–10.
15. Leekitcharoenphon P, Nielsen E, Kaas R, Lund O, Aarestrup F. Evaluation of Whole Genome Sequencing for Outbreak Detection of *Salmonella enterica*. *PLoS ONE*. 2014;9(2):e87991.
16. Grad Y, Lipsitch M, Feldgarden M, Arachchi H, Cerqueira G, Fitzgerald M, et al. Genomic epidemiology of the *Escherichia coli* O104:H4 outbreaks in Europe, 2011. *Proc Natl Acad Sci U S A*. 2011;109(8):3065–70.
17. McDonnell J, Dallman T, Atkin S, Turbitt D, Connor T, Grant K, et al. Retrospective analysis of whole genome sequencing compared to prospective typing data in further informing the epidemiological investigation of an outbreak of *Shigella sonnei* in the UK. *Epidemiol Infect*. 2013;141:2568–75.
18. Underwood A, Dallman T, Thomson N, Williams M, Harker K, Perry N, et al. Public health value of next-generation DNA sequencing of enterohemorrhagic *Escherichia coli* isolates from an outbreak. *J Clin Microbiol*. 2013;51(1):232–7.
19. Lavezzo E, Toppo S, Franchin E, Di Camillo B, Finotello F, Falda M et al. Genomic comparative analysis and gene function prediction in infectious diseases: application to the investigation of a meningitis outbreak. *BMC Infect Dis*. 2013;13(554). doi:10.1186/1471-2334-13-554.
20. Miller R, Price J, Batty E, Didelot X, Wyllie D, Golubchik T, et al. Healthcare-associated outbreak of methicillin-resistant *Staphylococcus aureus* bacteraemia: role of a cryptic variant of an epidemic clone. *J Hosp Infect*. 2014;86(2):83–9.
21. Hawkey J, Edwards D, Dimovski K, Hiley L, Billman-Jacobe H, Hogg G et al. Evidence of microevolution of *Salmonella* Typhimurium during a series of egg-associated outbreaks linked to a single chicken farm. *BMC Genomics*. 2013;14(800). doi:10.1186/1471-2164-14-800.
22. Larsson J, Torpdahl M, Petersen R, Sørensen G, Lindsted B, Nielsen E. Development of a new nomenclature for *Salmonella* Typhimurium multilocus variable number of tandem repeats analysis (MLVA). *Euro Surveill*. 2009;14(15).
23. Wang Q, Kong F, Jelfs P, Gilbert G. Extended phage locus typing of *Salmonella enterica* serovar Typhimurium, using multiplex PCR-based reverse line blot hybridization. *J Med Microbiol*. 2008;57:827–38.
24. Lindstedt B, Vardund T, Aas L, Kapperud G. Multiple-locus variable-number tandem-repeats analysis of *Salmonella enterica* subsp. *enterica* serovar Typhimurium using PCR multiplexing and multicolor capillary electrophoresis. *J Microbiol Methods*. 2004;59:163–72.
25. Octavia S, Wang Q, Tanaka M, Kaur S, Sintchenko V, Lan R. Delineating Community Outbreaks of *Salmonella enterica* Serovar Typhimurium by Use of Whole-Genome Sequencing: Insights into Genomic Variability within an Outbreak. *J Clin Microbiol*. 2015;53(4):1063–71.
26. Dimovski K, Cao H, Wijburg O, Strugnelli R, Mantena R, Whipp M, et al. Analysis of *Salmonella enterica* Serovar Typhimurium variable-number tandem-repeat data for public health investigation based on measured mutation rates and whole-genome sequence comparisons. *J Bacteriol*. 2014;196(16):3036–44.
27. WHO Collaborating Centre for Reference and Research on Salmonella. Antigenic Formulae of the *Salmonella* Serovars 9th ed. Paris: World Health Organisation; 2007.
28. Fu S, Octavia S, Tanaka M, Sintchenko V, Lan R. Defining the core genome of *Salmonella enterica* serovar Typhimurium for genomic surveillance and epidemiological typing. *J Clin Microbiol*. 2015;53(8):2530–8.
29. Pang S, Octavia S, Feng L, Liu B, Reeves PR, Lan R, et al. Genomic diversity and adaptation of *Salmonella enterica* serovar Typhimurium from analysis of six genomes of different phage types. *BMC Genomics*. 2013;14:718.
30. Kingsley R, Msefula C, Thomson N, Kariuki S, Holt K, Gordon M, et al. Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Res*. 2009;19(12):2279–87.
31. Mather A, Reid S, Maskell D, Parkhill J, Fookes M, Harris S, et al. Distinguishable epidemics within different hosts of the multidrug resistant Zoonotic pathogen *Salmonella* Typhimurium DT104. *Science*. 2013;341(6153):1514–7.
32. Okoro C, Kingsley A, Connor T, Harris R, Parry C, Al-Mashhadani M, et al. Intra-continental spread of human invasive *Salmonella* Typhimurium pathovariants in sub-Saharan Africa. *Nat Genet*. 2012;44(11):1215–21.
33. Inns T, Lane C, Peters T, Dallman T, Chatt C, McFarland N, et al. A multicountry *Salmonella enteritidis* phage type 14B outbreak associated with eggs from a German producer: near real-time application of whole genome sequencing and food chain investigations, United Kingdom, May to September 2014. *Euro Surveill*. 2015;20(16):21098.
34. Byrne L, Fisher I, Peters T, Mather A, Thomson N, Rosner B, et al. A multi-country outbreak of *Salmonella* Newport gastroenteritis in Europe associated with watermelon from Brazil, confirmed by whole genome sequencing: October 2011 to January 2012. *Euro Surveill*. 2014;19(31):6–13.
35. den Bakker H, Allard M, Bopp D, Brown E, Fontana J, Iqbal Z, et al. Rapid whole-genome sequencing for surveillance of *Salmonella enterica* serovar Enteritidis. *Emerg Infect Dis*. 2014;20(8):1306–14.
36. Food US, Administration D. Whole Genome Sequencing (WGS) Program. Silver Spring: US Department of Health and Human Services; 2015. <http://www.fda.gov/Food/FoodScienceResearch/WholeGenomeSequencingProgramWGS/>. Accessed 3 May 2015.
37. Price J, Gordon N, Crook D, Llewelyn M, Paul J. The usefulness of whole genome sequencing in the management of *Staphylococcus aureus* infections. *Clin Microbiol Infect*. 2013;19(9):784–9.
38. Sabat A, Budimir A, Nashev D, Sa-Leao R, van Dji J, Laurent F, et al. Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro Surveill*. 2013;18(4):20380.
39. Cheung M, Kwan H. Fighting outbreaks with bacterial genomics: case review and workflow proposal. *Public Health Genomics*. 2012;15(6):341–51.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

