bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Full Title: The diversity, evolution and ecology of *Salmonella* in venomous snakes

Short Title: *Salmonella* diversity in venomous snakes

Caisey V. Pulford[1], Nicolas Wenner[1], Martha L. Redway[1,#a], Ella V. Rodwell[1], Hermione J. Webster[1], Roberta Escudero[1,#b], Carsten Kröger[1,#c], Rocío Canals[1], Will Rowe[1,#d], Javier Lopez[2], Neil Hall[3,4] Paul D Rowley[5], Dorina Timofte[6,7], Robert A. Harrison[5], Kate S. Baker[1], Jay C. D. Hinton[1*]

[1]Departement of Functional and Comparative Genomics, Institute of Integrative Biology, University of Liverpool, Liverpool, UK

[2]Chester Zoo, Cheshire, UK

[3]Earlham Institute, Norwich Research Park, Norwich, UK

[4]School of Biological Sciences, University of East Anglia, Norwich, UK

[5]Centre for Snakebite Research and Interventions, Liverpool School of Tropical Medicine, Liverpool, UK

[6]Institute of Veterinary Science, University of Liverpool, Leahurst Campus, Cheshire, UK

[7]Institute of Infection and Global Health, University of Liverpool, UK

[#a]Current address: Department of Human Nutrition, University of Glasgow, Glasgow, UK

[#b]Current address: Centro de Ciências da Saúde, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

[#c]Current address: Department of Microbiology, School of Genetics and Microbiology, Moyne Institute of Preventive Medicine, Trinity College Dublin, Dublin, Ireland

[#d]Current address: Science and Technology Facilities Council, Daresbury, UK

* Corresponding Author

Email: jay.hinton@liv.ac.uk (JCDH)

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Abstract

Background

Reptile-associated *Salmonella* are a major, but often neglected cause of both gastrointestinal and bloodstream infection globally. The diversity of *Salmonella enterica* has not yet been determined in venomous snakes, however other cold-blooded animals have been reported to carry a broad range of *Salmonella* bacteria. We investigated the prevalence and assortment of *Salmonella* in a collection of venomous snakes in comparison with non-venomous reptiles.

Methodology/Principle Findings

We used a combination of selective enrichment techniques and whole-genome sequencing. We established a unique dataset of reptilian isolates to study *Salmonella enterica* species-level evolution and ecology and investigated differences between phylogenetic groups. We observed that 91% of venomous snakes carried *Salmonella,* and found substantial diversity between the serovars (n=58) carried by reptiles. The *Salmonella* serovars belonged to four of the six *Salmonella enterica* subspecies: *diarizonae*, *enterica*, *houtanae* and *salamae.* Subspecies *enterica* isolates were distributed among two distinct phylogenetic clusters, previously described as clade A (52%) and clade B (48%). We identified metabolic differences between *S. diarizonae, S. enterica* clade A and clade B involving growth on lactose, tartaric acid, dulcitol, *myo*-inositol and allantoin.

Significance

We present the first whole genome-based comparative study of the *Salmonella* bacteria that colonise venomous and non-venomous reptiles and shed new light on *Salmonella* evolution. The findings raise the possibility that venomous snakes are a reservoir for human Salmonellosis in Africa. The proximity of venomous snakes to human dwellings in rural Africa may result in contaminated faecal matter being shed on surfaces and in water sources used

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

for human homes and to irrigate salad crops. Because most of the venomous snakes had been captured in Africa, we conclude that the high level of *Salmonella* diversity reflects the African environmental niches where the snakes have inhabited.

Author Summary

Salmonella enterica is a remarkable bacterial species that causes Neglected Tropical Diseases globally. The burden of disease is greatest in some of the most poverty-afflicted regions of Africa, where salmonellosis frequently causes bloodstream infection with fatal consequences. The bacteria have the ability to colonise the gastrointestinal tract of a wide range of animals including reptiles. Direct or indirect contact between reptiles and humans can cause Salmonellosis. In this study, we determined the prevalence and diversity of *Salmonella* in a collection of African venomous snakes for the first time. Using the power of genomics, we showed that the majority of venomous snakes (91%) carry *Salmonella*, two thirds of which belonged to a subspecies of *S. enterica* called *enterica,* which is associated with most cases of human salmonellosis. Within the *S. enterica* subspecies we identified two evolutionary groups which display distinct growth patterns on infection relevant carbon sources. Our findings could have particular significance in Africa where venomous snakes wander freely around human dwellings and potentially shed contaminated faecal matter in water sources and on surfaces in rural homes.

Introduction

Salmonella is a clinically relevant bacterial pathogen of great public health significance [1,2]. The *Salmonella* genus contains two species; *S. bongori* and *S. enterica. S. enterica* is further divided into six subspecies; *enterica* (I), *salamae* (II), *arizonae* (IIIa), *diarizonae* (IIIb), *houtanae* (IV) and *indicia* (VI) [3]. The subspecies are classified into approximately 2 600 serovars which are ecologically, phenotypically and genetically diverse [4]. Serovars which belong to *S. enterica* subspecies *enterica* cluster phylogenetically into two predominant

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

83 clades (A and B) [5–7]. Henceforth *Salmonella* will refer to the *S. enterica* species only,
84 unless stated otherwise.
85
86 At the genus level, *Salmonella* has a broad host-range whilst individual serovars differ in
87 host-specificity [8]. The majority of *Salmonella* infections in humans (99%) are caused by a
88 small number of serovars belonging to the *S. enterica* subspecies [9]. Serovars which
89 belong to non-*enterica* subspecies are associated with carriage in cold-blooded animals
90 such as non-venomous reptiles and amphibians, but are rarely found in humans [8,10–12].
91 Carriage rates of non-*enterica* serovars in reptiles can be high, for example, one study
92 focused on snakes in a pet shop found that 81% of animals were carrying *S. diarizonae* [12].
93 Further examples of previous studies demonstrating the diverse range of *Salmonella*
94 serovars that colonise various reptilian species in different countries are shown in Table 1.
95
96 **Table 1. The distribution of *Salmonella* subspecies in non-venomous reptiles**

Citation	Country of study	Reptile information	Distribution
Piasecki *et al.,* 2014 [13]	Poland	Snakes, Lizards, Chelonians	59% *enterica* 16% *salamae* or *houtanae* 3.5% *diarizonae* or *indica*
Lukac *et al.,* 2015 [14]	Croatia	Snakes, Lizards, Chelonians	34.6% *enterica* 23.1% *houtanae* 23.1% *arizonae* 15.4% *diarizonae* 2.8% *salamae*
Nakadai *et al.,* 2005 [15]	Japan	Snakes, Lizards, Turtles	62.5% *enterica* 37.5% *salamae, arizonae, diarizonae* or *houtanae*
Geue and Löschener, 2002 [16]	Germany and Austria	Reptiles	52.8% *enterica* 34.5% *diarizonae* 3.4% *salamae* 6.9% *arizonae* 2.3% *houtanae*
Sá and Solari, 2001 [17]	Brazil	Brazilian and Imported Pet Snakes, Lizards and Chelonians	44.7% *enterica* 10.5% *salamae* 5.2% *arizonae* 21% *diarizonae* 18.5% *houtanae*
Perderson *et al.,* 2009 [18]	Denmark	Captive reptiles	65% *enterica* 12% *diarizonae* 11% *houtanae* 6% *salamae* 6% *arizonae*
Schröter *et al.,* 2004 [12]	Germany	Captive reptiles	100% *diarizonae*

97

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Reptiles represent a significant reservoir for serovars of *Salmonella* associated with human disease, as over 60% of reptiles have been reported to carry serovars falling within the *S. enterica* subspecies [18]. About 6% of human salmonellosis cases are contracted from reptiles in the USA [19], and in South West England, 27.4% of *Salmonella* cases in children under five years old were linked to reptile exposure [20]. The latter study demonstrated that reptile-derived salmonellosis was more likely to cause bloodstream infection in humans than non-reptile-derived *Salmonella* [20]. Reptile-associated *Salmonella* is therefore considered to be a global threat to public health [21].

The majority of reptile-associated salmonellosis cases reported in humans are caused by *Salmonella* from non-venomous reptiles [21], most likely because these are more frequently kept as pets. Therefore, non-venomous reptiles have been the focus of numerous studies while the prevalence and diversity of *Salmonella* in venomous snakes remained unknown. The evolution of venom function is a key factor that drives ecological diversification in snakes [22], with venomous snakes differing from non-venomous reptiles in a number of important aspects including the way that animals interact with the environment, food sources, habitat and humans which may have concomitant effects upon *Salmonella* carriage.

Research to improve snakebite treatment at the Liverpool School of Tropical Medicine (LSTM) has resulted in the creation of the UK's most extensive collection of venomous snakes (195). The LSTM herpetarium houses venomous snakes from a diverse range of species and geographical origins representing an ideal source of samples to assess *Salmonella* in this under-studied ecological niche. We collected *Salmonella* isolates from 97 venomous snakes housed at LSTM, and made genome-level comparisons with *Salmonella* collected from 28 non-venomous reptiles from the University of Liverpool veterinary diagnostic laboratory.

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

The aims of this study were three-fold. Firstly, to determine the period prevalence of *Salmonella* in a collection of captive venomous snakes and investigate whether this group of reptiles are potential reservoirs for human salmonellosis in Africa. Secondly, to assess the serological and phylogenetic diversity of *Salmonella* amongst venomous snakes compared with non-venomous reptiles. Thirdly, to utilise the diversity of *Salmonella* carried in both venomous snakes and non-venomous reptiles to assess clade-specific differences thought to reflect adaptation to survival in the environment or different hosts. Here, we present the first whole genome-based comparative study of the *Salmonella* bacteria that colonise venomous and non-venomous reptiles.

Methods

Source of *Salmonella* isolates

The *Salmonella* isolates were derived from faecal samples from two collections of reptiles. One hundred and six faecal samples were collected from venomous snakes at LSTM from May 2015 to January 2017, with an emphasis on snakes originating from Africa (S1 Table), and investigated for the presence of *Salmonella.* All venomous snakes where housed in individual enclosures and fed with frozen mice. Sixty-nine of the samples (71%) were sourced from wild-caught snakes originating from: Togo, Nigeria, Cameroon, Egypt, Tanzania, Kenya, South Africa, and Uganda. A further 28 *Salmonella* isolates (29%) came from venomous snakes bred in captivity. The LSTM herpetarium is a UK Home Office licensed and inspected animal holding facility. A second collection of 28 *Salmonella* isolates from non-venomous reptiles were sourced from the veterinary diagnostics laboratory based at the University of Liverpool's Leahurst campus (reptilian species described in S1 Table). Leahurst isolate 11L-2493 was isolated from a beaded lizard which has the ability to produce venom. These isolates were collected from June 2011 to July 2016 from specimens submitted as part of *Salmonella* surveillance for import/export, but also clinical faecal samples and tissues from post mortem investigations. The provenance of the non-venomous

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

153 reptile isolates is described in S1 Table. The majority of the non-venomous reptiles were
154 sourced from a zoological collection, however two animals were privately owned and three
155 were sourced from the RSPCA. The LSTM isolates are henceforth referred to as venomous
156 snake isolates and the Leahurst isolates are referred to as non-venomous reptile isolates.
157

158 **Isolation of *Salmonella***

159 All media were prepared and used in accordance with the manufacturers guidelines unless
160 otherwise stated. *Salmonella* was isolated using a modified version of the protocol described
161 in the national Standard Operating Procedure for detection of *Salmonella* issued by Public
162 Health England [23].
163

164 Faecal droppings were collected from reptiles and stored in 15 mL plastic centrifuge tubes at
165 4˚C. Two different methods were used for the enrichment of *Salmonella* from faecal samples
166 due to reagent availability at the time of isolation. S1 Table provides information on isolate
167 specific methods. In enrichment method 1, faecal samples were added to 10 ml of buffered
168 peptone water (Fluka Analytical, UK, 08105-500G-F) and incubated overnight at 37˚C with
169 shaking at 220 rpm. Following overnight incubation, 100 µL of the faeces mixture was added
170 to 10 mL of Selenite Broth (19 g/L selenite broth base, Merck, UK, 70153-500G and 4 g/L
171 sodium hydrogen selenite, Merck 1.06340-50G) and incubated overnight at 37°C with
172 shaking at 220 rpm. In enrichment method 2, faecal samples were added to 10 ml of
173 Buffered Peptone Water (Fluka Analytical, 08105-500G-F) supplemented with 10 µg/ml
174 Novobiocin (Merck, N1628), and incubated overnight at 37˚C with shaking at 220 rpm.
175 Following overnight incubation, 100 µL of the faeces mixture was added to 10 ml Rappaport-
176 Vassilliadis Medium (Lab M, UK, LAB086) and incubated for 24 hours at 42˚C with shaking
177 at 220 rpm.
178

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Following enrichment, 10 µL of overnight broth was spread onto Xylose Lysine Deoxycholate (XLD) (Oxoid, UK, CM0469) agar plates which were incubated overnight at 37˚C. Putative *Salmonella* colonies were selected by black appearance on XLD plates and confirmed by pink and white colony formation on Brilliant Green Agar (Merck, 70134-500G) supplemented with 0.35 g/L mandelic acid (Merck, M2101) and 1 g/L sodium sulfacetamide (Merck, S8647).

To identify *S. enterica* species, colony PCR of the *Salmonella* specific *ttr* locus, which is required for tetrathionate respiration [24], was performed. PCR reagents included MyTaq Red Mix 1x (Bioline, UK, BIO-25043), *ttr-4* reverse primer (5'-AGCTCAGACCAAAAGTGACCATC-3') and *ttr-6* forward primer (5'-CTCACCAGGAGATTACAACATGG-3') on colonies suspected to be *Salmonella*. PCR reaction conditions were as follows: 95˚C 2 min, 35 x (95˚C 15 s, 60˚C 30 s, 72˚C 10 s), 72˚C 5 min. PCR products were visualised using agarose gel (3.5%) (Bioline, BIO-41025) electrophoresis in TAE buffer. Midori Green DNA stain (3 µL/100 mL) (Nippon Genetics, Germany, MG 04) was used to visualise DNA bands under UV light. Throughout the isolation procedure, *S. enterica* serovar Typhimurium (*S*. Typhimurium) strain LT2 [25,26] was used as a positive control, and *Escherichia coli* MG1655 [27] was used as a negative control (S1 Table).

Whole-genome sequencing of *Salmonella* from venomous and non-venomous reptiles

All non-venomous reptile isolates and 87 of 97 venomous snake isolates were sent for whole genome sequencing. LSS-88 to LSS-97 were not sequenced as they were collected after the submission deadline to the sequencing centre. Isolates were sent to either MicrobesNG, UK or the Earlham Institute, UK for whole-genome sequencing on the Illumina HiSeq platform (Illumina, California, USA). Isolates which were sequenced by MicrobesNG were

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

prepared for sequencing in accordance with the company's preparation protocol. Isolates which were sequenced by the Earlham Institute were prepared by inoculating a single colony of *Salmonella* into a FluidX® 2D Sequencing Tube (FluidX Ltd, UK) containing 100 µL of Lysogeny Broth (LB, Lennox) and incubating overnight at 37˚C, with shaking at 220 rpm. LB was made using 10 g/L bacto tryptone (BD Biosciences, UK, 211705), 5 g/L bacto yeast extract (BD, 212750) and 5 g/L sodium chloride (Merck, S3014-1kg). Following overnight growth, the FluidX® 2D Tubes were placed in a 95˚C oven for 20 minutes to heat-kill the isolates.

DNA extractions and Illumina library preparations were conducted using automated robots at MicrobesNG or the Earlham Institute. At the Earlham institute, the Illumina Nextera XT DNA Library Prep Kit (Illumina, FC-131-1096) was used for library preparation. High throughput sequencing was performed using an Illumina HiSeq 4 000 sequencing machine to achieve 150 bp paired-end reads. Sequencing was multiplexed to 768 unique barcode combinations per sequencing lane. The insert size was determined to be approximately 180 bp and the median depth of coverage was 30x.

At MicrobesNG, genomic DNA libraries were prepared using the Nextera XT Library Prep Kit (Illumina, FC-131-1096) with two nanograms of DNA used as input and double the elongation time of that described by the manufacturer. Libraries were sequenced on the Illumina HiSeq 2 500 using a 250 bp protocol.

Obtaining contextual reference sequences

Reads were uploaded to Enterobase for serovar prediction using the *Salmonella in silico* typing resource (SISTR) [28,29]. Enterobase was also used to assign a Multi Locus Sequence Type (MLST) to each isolate, based on sequence conservation of seven housekeeping genes [29]. Where available, reference isolates representing previously sequenced *Salmonella* isolates for all subspecies and serovars identified were included in

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

234 the analysis. Reference sequence assemblies were downloaded from the National Center
235 for Biotechnology Information (NCBI). Accession numbers are available in S2 Table.
236

237 **Quality control checks**

238 Fastqc v0.11.5 (www.bioinformatics.babraham.ac.uk/projects/fastqc/) and multiqc v1.0
239 (http://multiqc.info) were used to assess read quality. Kraken v0.10.5-beta [30] was run to
240 ensure reads were free from contamination using the MiniKraken 8gb database and a
241 *Salmonella* abundance cut-off of 70%. Trimmomatic v0.36 [31] was then used on the paired
242 end reads to trim low-quality regions using a sliding widow of 4:200. ILLUMINACLIP was
243 used to remove adapter sequences.
244

245 **Phylogenetics- core genome alignment tree**

246 Genomes were assembled using SPADES v3.90 [32]. QUAST v4.6.3 [33] was used to
247 assess the quality of assemblies, the results of which can be found in S3 Table. Assemblies
248 which comprised of greater than 500 contiguous sequences were deemed too fragmented
249 for downstream analysis. All assemblies which passed QC were annotated using Prokka
250 v1.12 [34]. Roary v3.11.0 [35] was used to generate a core genome alignment. SNP-sites
251 v2.3.3 [36] was used to extract SNPs. A maximum likelihood tree was built from the core
252 genome SNP alignment of all isolates using RAxML-NG v0.4.1 BETA [37] with the general
253 time reversible GTR model and gamma distribution for site specific variation and 100
254 bootstrap replicates to assess support. The tree was rooted using the *Salmonella* species *S.*
255 *bongori.* Interactive Tree Of Life v4.2 [38] was used for tree visualisation. We confirmed that
256 there was no bias in phylogenetic signal between the two different sequencing platforms
257 used by assessing clustering patterns within the phylogenies. S1 Table contains details of
258 the sequencing facility. Monophyletic clustering of isolates was used to assign subspecies to
259 newly sequenced *Salmonella* from venomous and non-venomous reptilian hosts. The level

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

of association between venom status and phylogenetic clade was determined using odds ratios and χ2 statistics using the OpenEpi website (http://www.openepi.com).

Identification of clade-specific genomic regions

Genes involved in the utilisation of each carbon source were identified using KEGG [39]. Sequences were downloaded using the online tool SalComMac [40], which allows the download of fasta sequences of the genes in *S.* Typhimurium strain 4/74. In the case of *lac* genes the sequences were taken from the *E. coli* reference sequence MG1655. The sequences can be found in S1 Text.

The software tool MEGABLAST v2.2.17 [41] was used to perform a BLAST search of genes in the reptile-derived genomes against a custom-made database of genes diagnostic of *Salmonella* Pathogenicity Islands and genes involved in carbon utilisation. To confirm all MEGABLAST results, the short reads were mapped against each gene using BWA v0.7.10 [42] and SAMtools v0.1.19 [43]. The resulting bam files were manually assessed for gene presence and absence using Integrative Genomics Viewer v2.4.15 [44]. The results were plotted against the maximum likelihood phylogeny using Interactive Tree Of Life v4.2 [38].

Carbon source utilisation

Differential carbon source utilisation of 39 reptile-derived *Salmonella* isolates from *S. diarizonae, S. enterica* clade A and *S. enterica* clade B was assessed. Filter-sterilised carbon sugar solutions were added into M9 (Merck, M6030-1kg) agar at concentrations which are described in S4 Table. Isolated colonies were transferred from LB agar plates onto M9 carbon source plates using a sterile 48-pronged replica plate stamp and incubated at 37°C firstly under aerobic conditions. If no growth was seen under aerobic conditions for a carbon source, then the procedure was repeated under anaerobic conditions (approx. 0.35% oxygen) with 20 mM Trimethylamine *N*-oxide dehydrate (TMAO) (Merck, 92277) as a

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

terminal electron acceptor. Anaerobic conditions were achieved by incubating plates in an anaerobic jar with 3x AnaeroGen 2.5 L sachets (Thermo Scientific, UK, AN0025A) to generate anaerobic gas. Oxygen levels were measured using SP-PSt7-NAU Sensor Spots and the Microx 4 oxygen detection system (PreSens, Regensburg, Germany). An LB control plate was used to validate successful bacterial transfer and all experiments were performed in duplicate. *Salmonella* growth was determined at 18, 90 and 162 hours in aerobic growth conditions and at 162 hours in anaerobic growth conditions. A sub-set of growth positive isolates were assessed for single colony formation to validate the results of the replica plating.

Antimicrobial susceptibility testing of *Salmonella* isolated from venomous and non-venomous reptiles

Antimicrobial susceptibility was determined using a modified version of the European Committee on Antimicrobial Susceptibility Testing (EUCAST) disk diffusion method [45] using Mueller Hinton (Lab M, LAB039) agar plates and a DISKMASTER ™ dispenser (Mast Group, UK, MDD64). Inhibition zone diameters were measured and compared to EUCAST zone diameter breakpoints for Enterobacteriaceae [46]. Isolates were first tested with six commonly used antibiotics (Ampicillin 10 µg, Chloramphenicol 30 µg, Nalidixic Acid 30 µg, Tetracycline 30 µg, Ceftriaxone 30 µg, and Trimethoprim/Sulfamethoxazole 25 µg) and were then tested with five additional antibiotics (Meropenem 10 µg, Gentamicin 10 µg, Amoxicillin/Clavulanic Acid 30 µg, Azithromycin 15 µg, and Ciprofloxacin 5 µg) if any resistance was seen to the primary antibiotics (all disks from Mast Group). If resistance was observed phenotypically, then the presence of antimicrobial resistance genes were investigated using the Resfinder software tool [47]. Antimicrobial resistance was defined as resistance to one antimicrobial to which isolates would normally be susceptible [48]. Multidrug resistance was defined as an isolate which showed resistance to three or more antimicrobials to which it would normally be susceptible [48].

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

314

315 **Data availability**

316 Fastq files are available at NCBI Short Read Archive under the accession numbers listed in
317 S1 Table (BioProject Number PRJNA506152, SRA Study Number SRP170008.)

318

319 # Results and Discussion

320 ## *Salmonella* prevalence in venomous snakes is similar to that identified in non-
321 venomous reptiles

322 *Salmonella* carriage is well documented amongst reptiles (Table 1), however, to our
323 knowledge there is no published study that reports on *Salmonella* in venomous snakes. The
324 period prevalence of *Salmonella* was assessed in a collection of 106 venomous snakes
325 housed at the LSTM venom unit between May 2015 and January 2017. A remarkably high
326 proportion (91%; 97/106) of the faecal samples contained *Salmonella* (S1 Table), which
327 should be seen in the context of the significant carriage rate of *Salmonella* by other non-
328 venomous reptiles described in the literature [21]. Our findings pose important public health
329 considerations for individuals who work with venomous snakes, which may previously have
330 been overlooked.

331

332 ## *Salmonella* diversity in venomous snakes and non-venomous reptiles
333 highlights the possibility of local transmission events and long-term shedding

334 To assess diversity, 87 venomous snake-derived *Salmonella* isolates and 28 non-venomous
335 reptile-derived *Salmonella* isolates were whole genome sequenced. *In silico* serotyping
336 revealed 58 different *Salmonella* serovars (Fig 1). Multi-locus sequence typing was used to
337 perform sub-serovar genetic characterisation of *Salmonella* [29]. Isolates falling within the
338 same serovar had identical sequence types (S1 Table), reflecting the intra-serovar
339 homogeneity of the *Salmonella* isolated in this study. Taking into account the total number of
340 samples assessed, serovar diversity was greater in non-venomous reptiles (n of serovars =

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

22, n of samples = 28) compared to venomous snakes (n of serovars = 38, n of samples = 87).

Fig 1. The distribution of the 58 *Salmonella enterica* serovars isolated from venomous and non-venomous reptiles. Each bar represents the total number of isolates which belonged to each serovar. Serovars containing isolates that had multiple serovar designations from SISTR are indicated with asterisks. Human pictographs are displayed on serovars which are amongst the top 20 isolated from humans in Africa. Data is based on the global monitoring of *Salmonella* serovar distribution from the WHO global foodborne infections network data [49].

The most common serovar to be identified in venomous snakes was *S.* Souhanina (n = 15). All *S.* Souhanina isolates cluster locally on the phylogeny (Fig 2) falling within a 5 SNP cluster characteristic of a clonal expansion event [50]. Four of these isolates were found in captive bred reptiles, whilst 11 isolates came from venomous snakes which originated in Cameroon, Uganda, Tanzania, Nigeria, Togo and Egypt. The close phylogenetic relationship between *S.* Souhanina isolates which belong to the same MLST type (ST488) from captive and imported animals with a range of origins suggests that some local *Salmonella* transmission may be occurring. Transmission of *Salmonella* between captured reptiles has been associated with consumption of contaminated food sources such as frozen mice [51,52]. However, a single food source would not explain the high prevalence of venomous snakes which carried distinct and diverse serovars and MLST types.

Fig 2. The diversity of *Salmonella* isolated from a collection of venomous snakes and non-venomous reptiles. Core genome maximum likelihood phylogenetic tree. Tree was rooted using *S. bongori* – not shown (S1 Figure). 25 contextual reference genomes representing previously sequenced isolates from each *Salmonella* subgroup are indicated in red. Colour strips showing metadata are as follows; Subspecies – Subspecies of isolate,

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Origin – The country of origin of the snake from which the isolate was taken, V or NV – Depicts whether the reptile host was venomous (V) or non-venomous (NV), AMR phenotype- Isolates shown in red were resistant to one or more antimicrobial agent. Metadata for the contextual reference genomes appear as white. Tree was visualised using ITOL (https://itol.embl.de).

As long term intermittent shedding of *Salmonella* from reptiles has previously been described [54], we assessed the continuity of *Salmonella* shedding from venomous snakes, by collecting three faecal samples from a Western Green Mamba from Togo over a three-month period between 31st October 2016 and 31st January 2017. All three faecal samples contained *Salmonella* which belonged to sequence type ST488, showing that individual snakes can shed the same sequence type of *Salmonella* over a reasonably long period of time. This continuous shedding of *Salmonella* by a venomous snake demonstrates the risk that reptiles pose to public health. Many of the snakes had been maintained in the LSTM herpetarium for years after collection from the wild. The diverse range of serovars which were isolated suggests that the venomous snakes in this study acquired *Salmonella* in Africa, prior to entering the unit and can shed the bacteria for many years.

Venomous reptiles carry multidrug resistant *Salmonella* serovars of clinical relevance

Because the majority of venomous snakes examined in this study were of African origin or belonged to a species of snake native to the African continent (Fig 2), we compared the *Salmonella* serovars isolated from all reptiles in this study with those most frequently associated with causing human disease in Africa. The *Salmonella* serovar distribution has been reported by the WHO global foodborne infections network data bank based on data from quality assured laboratories in Cameroon, Senegal and Tunisia [49] (Fig 1). Eleven isolates belonged to serovars commonly pathogenic in humans. We therefore determined

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

396 the proportion of all venomous snakes and non-venomous reptiles that carried antimicrobial
397 resistant *Salmonella* (Table 2). In *Salmonella* collected from venomous snakes, 4.1% of
398 isolates (4/97) were resistant to at least one antimicrobial and two isolates were multidrug
399 resistant (Table 2). Three resistant isolates from venomous snakes belonged to the serovar
400 Enteritidis and were closely related to the global *S.* Enteritidis epidemic clade which has
401 been shown to be associated with causing human disease in Africa [55]. One of the *S.*
402 Enteritidis isolates (LSS-32) was isolated from a snake which had been imported from
403 Nigeria. These findings raise the possibility that venomous snakes could spread *Salmonella*
404 in African settings where snakes wander freely around the wild and often inhabit rural
405 homes, particularly during the night [56–58]. Such speculation would require in-field
406 epidemiological validation. However, working alongside wild venomous snakes does present
407 logistical challenges.
408

409 **Table 2. Relating antimicrobial resistance to phenotype and genotype**

Isolate	Serovar[#]	Snake	Origin	AMR Resistance Phenotype*	Resistance Genes and Mutations	Antibiotic Family
LSS-6	Enteritidis	Monocled cobra	Captive Bred	Ampicillin Nalidixic Acid	*blaTEM-1B* *gyrA* p.D87Y (GAC to TAC) *aph(3")-1b* *sul2*	Beta-lactam Quinolones Aminoglycoside Sulphonamide
LSS-9	Enteritidis	Monocled cobra	Captive Bred	Ampicillin Chloramphenicol Nalidixic Acid Tetracycline Cotrimoxazole Azithromycin	*blaTEM-1B* *catA2*** *gyrA* p.D87G (GAC to GGC) *tet(A)* *sul2* *aph(3")-1b* *aph(6)-1d***	Beta-lactam Phenicol Quinolones Tetracycline Sulphonamide Aminoglycoside
LSS-28	Stanley	Malaysian Spitting Cobra	Captive Bred	Ampicillin Tetracycline	*blaOXA-1* *tet(A)*** *aph(4)-Ia* *aac(6')Ib-cr* *aac(3_1Va)*** *aac(6')Ib-cr* *sul2* *sul1* *catB3*** *ARR-3*	Beta-lactam Tetracycline Aminoglycoside Fluoroquinolone Sulphonamide Phenicol Rifampicin
LSS-32	Enteritidis	Black-Necked Spitting Cobra	Nigeria	Ampicillin Chloramphenicol Nalidixic Acid Tetracycline	*blaTEM-1B* *catA2*** *gyrA* p.D87G (GAC to GGC) *tet(A)* *aph(3")-1b* *aph(6)-1d*** *sul2*	Beta-lactam Phenicol Quinolones Tetracycline Aminoglycoside Sulphonamide
11L-2326	Typhimurium	Python	Captive Bred	Ampicillin Chloramphenicol	*blaCARB-2* *floR*	Beta-lactam Phenicol

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

				Nalidixic Acid Tetracycline Cotrimoxazole Azithromycin	*gyrA* p.D87Y (GAC to TAC) *tet(G)* *sul1* *aadA2*	Quinolones Tetracycline Sulfonamide Aminoglycoside
14L-1580	Unknown	Carpet Python	Captive Bred	Nalidixic Acid Tetracycline	*gyrA* p.D87G (GAC to GGC) *Tet(A)*	Quinolones Tetracycline
15L-2030	II 42:z:1,5	Dasia Olivacea	Captive Bred	Tetracycline	Non-identified genotypically on ResFinder	

Serovar predicted from genome sequence with SISTR [28]
* Determined experimentally, as described in Materials and Methods
** Less than 100% sequence identity on ResFinder

Phylogenetic diversity and molecular epidemiology of *Salmonella* in venomous snakes and non-venomous reptiles demonstrated for the first time

By studying *Salmonella* isolated from venomous and non-venomous reptiles we have shown that venomous snakes can shed *Salmonella.* The vast diversity of *Salmonella* has long been acknowledged in the literature [29], however it is rare to see such a broad representation of serovars in a single ecological niche. To study the diversity of reptile-associated *Salmonella* from an evolutionary perspective, we obtained 87 high quality whole genome sequences for the phylogenetic comparison with 24 contextual *Salmonella* genomes (methods). These included 60 isolates from venomous snakes, which were compared with 27 *Salmonella* isolates from non-venomous reptiles following a comprehensive comparative genomic analysis. We identified a total of 405 231 core genome SNPs that differentiated the 87 isolates, and were used to infer a maximum likelihood phylogeny (Fig 2). SNPs are considered to be a valuable marker of genetic diversity [50], and the hundreds of thousands of core-genome SNPs reflect a high level of genetic diversity between the reptile associated *Salmonella* isolates. The collection of reptile-derived *Salmonella* represented most of the known diversity of the *Salmonella* genus [3], spanning four of the six *Salmonella enterica* subspecies: *diarizonae*, *enterica*, *houtanae* and *salamae.* Reptile-derived *S. enterica* subspecies *enterica* isolates were approximately equally distributed among into two distinct phylogenetic clusters, known as clade A (58%) and clade B (48%) [5–7] (Fig 2). No

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

433 significant association was found between venom status and phylogenetic group (OR = 1.1,
434 CI = 0.3-3.0, $\chi 2$ = 0.02, P = 0.4).
435

436 **Genotypic and phenotypic conservation of infection-relevant carbon source**
437 **utilisation and virulence associated genomic regions sheds new light on**
438 ***Salmonella* ecology**

439 The unique collection of diverse *Salmonella* isolates was used to determine the phenotypic
440 and genotypic conservation of infection-relevant properties and genomic elements. Whilst
441 the reptile-associated *Salmonella* belonged to five evolutionary groups, the majority of
442 isolates were classified as *S. diarizonae* or *S. enterica*. The clustering of *S. enterica* into two
443 clades (A and B) has previously been inferred phylogenetically based on the alignment of 92
444 core loci [5,6]. The biological significance of *S. enterica* clade A and clade B has been
445 established as the two clades differ in host specificity, virulence-associated genes and
446 metabolic properties such as carbon utilisation [7]. The genome sequences were used to
447 expand upon pre-existing knowledge and determine phenotypic and genotypic conservation
448 of metabolic and virulence factors across *S. diarizonae* and *S. enterica* (clades A and B).
449

450 Although the majority of *Salmonella* serovars of public health significance belong to clade A,
451 certain clade B serovars such as *Salmonella* Panama have been associated with invasive
452 disease [59,60]. Clade B *S. enterica* generally carry a combination of two *Salmonella*
453 genomic islands: *Salmonella* Pathogenicity Island-18 and the cytolethal distending toxin
454 islet. It has been suggested that the two islands are associated with invasive disease, as
455 previously they had only been identified in *S. enterica* serovar Typhi and Paratyphi A, which
456 cause bloodstream infections [1,5]. The combination of *hylE* and *cdtB* genes were present
457 in all *S. diarizonae* and *S. enterica* clade B isolates in this study, but absent from all but one
458 *S. enterica* clade A isolates (14L-2174). We propose that the significant proportion of reptiles
459 which carried *S. enterica* clade B could partially explain the increased likelihood of snake-

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

460 associated salmonellosis causing invasive disease, compared to non-snake-acquired
461 salmonellosis [20].
462
463 The ability of *Salmonella* to grow in a wide range of conditions reflects the adaptation of the
464 bacteria to survival in the environment or in different hosts, as demonstrated by a recent
465 study focused on genome-scale metabolic models for 410 *Salmonella* isolates spanning 64
466 serovars in 530 different growth conditions [61]. We were interested in examining the
467 conservation of serovar-specific metabolic properties amongst the subspecies and clades of
468 *Salmonella* identified here, and the extent by which genotype reflected growth phenotype.
469 To assess metabolic differences between *S. enterica* clade A, *S. enterica* clade B and *S.*
470 *diarizonae*, we phenotypically screened 39 reptile isolates for the ability to catabolise a
471 number of infection-relevant carbon sources [55,62–64] (S4 Table). A summary of the
472 results for phenotypic carbon utilisation and the presence of genes associated with the
473 cognate metabolic pathway is shown in Fig 3.
474
475 **Fig 3. Phylogenetic context of carbon-source utilisation by reptile-derived-*Salmonella***
476 **isolates**. Carbon source data were mapped against the core genome phylogenetic tree. The
477 maximum likelihood tree includes all snake-derived *Salmonella* isolates which were
478 assessed for carbon utilisation and had high quality genome sequences. Reference
479 sequences for the majority of carbon genes were taken from *S.* Typhimurium strain 4/74, *lac*
480 gene sequences were taken from *E. coli* MG1655. Carbon sources which required anaerobic
481 conditions are indicated with asterisk.
482
483 In general, genotype accurately reflected phenotype in terms of carbon source utilisation;
484 however, this was not always the case (Fig 3). Discrepancies between phenotypic growth
485 and genotype suggests that mechanisms of *Salmonella* metabolism remain to be described.
486 For example, *S. diarizonae* isolate LSS-18 grew well on *myo*-inositol as a sole carbon
487 source (Fig 3) but showed zero percent homology with any *iol* genes present in the well-

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

488 characterised *Salmonella* strain 4/74. The ability to utilise lactose was a property of most *S.*
489 *diarizonae* isolates, consistent with previous reports that 85% of *S. diarizonae* are *Lac*$^{+}$ [65].
490 It is estimated that less than 1% of all *Salmonella* ferment lactose due to the loss of the *lac*
491 operon from the *S. enterica* subspecies [66]. It was interesting to discover that one non-
492 venomous snake isolate (13L-2837) which belongs to *S. enterica* clade B was capable of
493 utilising lactose as a sole carbon source. Isolate 13L-2337 belongs to the serovar *S.*
494 Johannesburg and to our knowledge this is the first published occurrence of a Lac^{+} *S.*
495 Johannesburg isolate. The 13L-2837 genome had zero percent homology to any *lac* genes
496 that corresponded to reference strain *E. coli* MG1655 (sequence in S1 Text) (results in Fig
497 3), suggesting an alternative method for lactose utilisation. The 13L-2837 *S.* Johannesburg
498 isolate also lacked the ability to grow on dulcitol, despite possessing all of the relevant *gat*
499 genes, raising the possibility of an inverse relationship between the ability of *Salmonella* to
500 utilise dulcitol and lactose as a sole carbon source.
501
502 The majority of *S. enterica* isolates from both clades utilised dulcitol, whereas dulcitol was
503 rarely used as a sole carbon source by *S. diarizonae*. These findings are consistent with
504 those of a study in Australian sleepy lizards, which demonstrated that dulcitol utilisation was
505 observed in almost all *S. enterica* and *S. salamae* isolates but only 10% of *S. diarizonae*
506 isolates [67]. Over 50% of the *S. enterica* clade A isolates lacked the *gatY* gene but grew
507 well on dulcitol as a sole carbon source. The *gatY* gene does not appear to be vital for
508 dulcitol catabolism. Differential repertoires of dulcitol catabolic genes across *Salmonella*
509 serovars are described by Nolle *et al.* (2017) [68], who found that *Salmonella* carries one of
510 two *gat* gene clusters. However, both of these clusters carry the *gatY* gene. Our findings
511 may indicate the possibility that a third *gat* gene cluster, is carried by some *Salmonella*
512 serovars.
513
514 In the majority of cases, allantoin was only utilised as a sole carbon source by *S. enterica*
515 clade A isolates, consistent with a previous report that described an association of clade A

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

516 with the allantoin catabolism island [5]. The ability of *Salmonella* to utilise allantoin is
517 considered to be an adaptation to mammalian and avian hosts, as allantoin is the end
518 product of the purine catabolic pathway in most mammals (except humans) and can be
519 found in the serum of birds [5,69]. In reptiles, the end product of the purine catabolic
520 pathway is not allantoin, but uric acid [5]. Absence of allantoin in the snake gastrointestinal
521 tract could explain why a substantial number of *S. enterica* clade B were found in snakes,
522 and supports the idea that clade B *Salmonella* are predominantly found in non-allantoin
523 producing hosts [70]. However, we identified one clade B isolate as an exception, isolate
524 11L-2351 which was sampled from a non-venomous reptile. This isolate belongs to the
525 serovar Montevideo, which is frequently associated with outbreaks of human salmonellosis
526 and shows a broad host range, unlike the other clade B serovars [71–73].
527
528 It is possible that the gain and loss of the allantoin catabolism genes may provide an
529 indication of host specificity. The relationship between the pseudogenization of the allantoin
530 metabolic genes and niche adaptation has also been proposed for the invasive non-
531 Typhoidal *Salmonella* (iNTS) reference isolate for *S.* Typhimurium: D23580 [74,75].
532 Compared with *S.* Typhimurium isolate 4/74, which shows a broad host range, D23580 is
533 unable to utilise allantoin as a carbon source, consistent with the adaptation of invasive
534 *Salmonella* in Africa towards non-allantoin producing hosts [74,75]. Furthermore, the
535 presence of pseudogene accumulation has been described in the allantoin degradation
536 pathway in host-restricted *Salmonella* serovars which cause invasive disease. Thus the loss
537 of ability to grow on allantoin is considered to be a marker of a switch from enteric to
538 invasive disease [69]. These findings may reflect the clinical observation that snake-acquired
539 *Salmonellosis* causes a more invasive disease that commonly results in hospitalisation,
540 compared to non-allantoin producing hosts [20].
541

542 ## Conclusion

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Reptiles are known to harbour a diverse range of *Salmonella* bacteria, however *Salmonella* carriage in a substantial proportion of reptile species has never been investigated. Here we have shown that venomous snakes harbour and shed a medically significant range of *Salmonella* serovars that represents much of the spectrum of the *Salmonella* genus and are phylogenetically distributed in a similar way to *Salmonella* found in non-venomous reptiles. Furthermore, we have identified that venomous snakes epidemiologically-linked to the African continent carry and excrete *Salmonella* serovars which cause human disease in Africa. In one case, the *Salmonella* isolate was resistant to first line antimicrobial agents. It is likely that venomous snakes represent a previously uncharacterised reservoir for *Salmonella* both in captive settings and in the African environment.

Furthermore, we have shown that reptiles are an ideal population of animals for the study of genus-level evolution of *Salmonella.* Reptiles represent a single ecological niche, which carry phylogenetically diverse isolates from the majority of *Salmonella* subspecies. By demonstrating the phenotypic and genotypic conservation of metabolic properties across three phylogenetic groups of *Salmonella* we have shed new light on the evolution of *Salmonella* serotypes.

Acknowledgements

We are grateful to present and former members of the Hinton lab and Lab H at the Institute of Integrative Biology for helpful discussions. In particular we would like to thank Alex V Predeus and Evelien M Adriaenssens for their ongoing bioinformatic advice and support. We are also grateful to Blanca Perez-Sepulveda for managing the 10 000 *Salmonella* genomes project, without whom sequencing of these isolates would not have been possible. We thank the Centre for Snakebite Research and Interventions at the Liverpool School of Tropical Medicine for access to faecal samples from venomous snakes and to the Leahurst

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

diagnostics unit at the University of Liverpool for providing access to the collection of *Salmonella* isolates from reptiles housed at Chester Zoo.

References

1. Crump JA, Luby SP, Mintz ED. The global burden of typhoid fever. Bull World Health Organ. Switzerland; 2004;82: 346–353. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2622843/
2. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, et al. The global burden of nontyphoidal *Salmonella* gastroenteritis. Clin Infect Dis. United States; 2010;50: 882–889. doi:10.1086/650733
3. Brenner FW, Villar RG, Angulo FJ, Tauxe R, Swaminathan B. *Salmonella* nomenclature. J Clin Microbiol. United States; 2000;38: 2465–2467. Available: http://jcm.asm.org/content/38/7/2465.full
4. Winfield MD, Groisman EA. Evolution and Ecology of *Salmonella*. EcoSal Plus. United States; 2004;1. doi:10.1128/ecosalplus.6.4.6
5. den Bakker HC, Moreno Switt AI, Govoni G, Cummings CA, Ranieri ML, Degoricija L, et al. Genome sequencing reveals diversification of virulence factor content and possible host adaptation in distinct subpopulations of *Salmonella enterica.* BMC Genomics. England; 2011;12: 425. doi:10.1186/1471-2164-12-425
6. Falush D, Torpdahl M, Didelot X, Conrad DF, Wilson DJ, Achtman M. Mismatch induced speciation in *Salmonella*: model and data. Philos Trans R Soc Lond B Biol Sci. England; 2006;361: 2045–2053. doi:10.1098/rstb.2006.1925
7. Parsons SK, Bull CM, Gordon DM. Substructure within *Salmonella enterica* subsp. *enterica* isolates from Australian wildlife. Appl Environ Microbiol. United States; 2011;77: 3151–3153. doi:10.1128/AEM.02764-10
8. Baumler A, Fang FC. Host specificity of bacterial pathogens. Cold Spring Harb Perspect Med. United States; 2013;3: a010041. doi:10.1101/cshperspect.a010041

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

596 9. Eng S-K, Pusparajah P, Ab Mutalib N-S, Ser H-L, Chan K-G, Lee L-H. *Salmonella*: A
597 review on pathogenesis, epidemiology and antibiotic resistance. Front Life Sci. Taylor
598 & Francis; 2015;8: 284–293. doi:10.1080/21553769.2015.1051243
599 10. Clancy MM, Davis M, Valitutto MT, Nelson K, Sykes JM 4th. *Salmonella* infection and
600 carriage in reptiles in a zoological collection. J Am Vet Med Assoc. United States;
601 2016;248: 1050–1059. doi:10.2460/javma.248.9.1050
602 11. Nair S, Wain J, Connell S, de Pinna E, Peters T. *Salmonella enterica* subspecies II
603 infections in England and Wales--the use of multilocus sequence typing to assist
604 serovar identification. J Med Microbiol. England; 2014;63: 831–834.
605 doi:10.1099/jmm.0.072702-0
606 12. Schröter M, Roggentin P, Hofmann J, Speicher A, Laufs R, Mack D. Pet snakes as a
607 reservoir for *Salmonella enterica* subsp. *diarizonae* (Serogroup IIIb): a prospective
608 study. Appl Environ Microbiol. United States; 2004;70: 613–615.
609 13. Piasecki T, Chrząstek K, Wieliczko A. *Salmonella* serovar spectrum associated with
610 reptiles in Poland. Acta Vet Brno. 2014;83: 287–294. Available:
611 https://doi.org/10.2754/avb201483040287
612 14. Lukac M, Pedersen K, Prukner-Radovcic E. Prevalence of *Salmonella* in captive
613 reptiles from Croatia. J Zoo Wildl Med. United States; 2015;46: 234–240.
614 doi:10.1638/2014-0098R1.1
615 15. Nakadai A, Kuroki T, Kato Y, Suzuki R, Yamai S, Yaginuma C, et al. Prevalence of
616 *Salmonella* spp. in pet reptiles in Japan. J Vet Med Sci. Japan; 2005;67: 97–101.
617 Available: https://www.jstage.jst.go.jp/article/jvms/67/1/67_1_97/_article
618 16. Geue L, Loschner U. *Salmonella enterica* in reptiles of German and Austrian origin.
619 Vet Microbiol. Netherlands; 2002;84: 79–91. Available:
620 https://www.sciencedirect.com/science/article/pii/S0378113501004370?via%3Dihub
621 17. Sá IVA de, Solari CA. *Salmonella* in Brazilian and imported pet reptiles [Internet].
622 Brazilian Journal of Microbiology. scielo; 2001. pp. 293–297. Available:
623 http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1517-

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

83822001000400007&lng=en&nrm=iso&tlng=en

18. Pedersen K, Lassen-Nielsen A-M, Nordentoft S, Hammer AS. Serovars of *Salmonella* from captive reptiles. Zoonoses Public Health. Germany; 2009;56: 238–242. doi:10.1111/j.1863-2378.2008.01196.x

19. Mermin J, Hutwagner L, Vugia D, Shallow S, Daily P, Bender J, et al. Reptiles, amphibians, and human *Salmonella* infection: a population-based, case-control study. Clin Infect Dis. United States; 2004;38 Suppl 3: S253-61. doi:10.1086/381594

20. Murphy D, Oshin F. Reptile-associated salmonellosis in children aged under 5 years in South West England. Arch Dis Child. England; 2015;100: 364–365. doi:10.1136/archdischild-2014-306134

21. Whiley H, Gardner MG, Ross K. A Review of Salmonella and Squamates (Lizards, Snakes and Amphibians): Implications for Public Health. Pathog (Basel, Switzerland). Switzerland; 2017;6. doi:10.3390/pathogens6030038

22. Fry BG, Vidal N, Norman JA, Vonk FJ, Scheib H, Ramjan SFR, et al. Early evolution of the venom system in lizards and snakes. Nature. Nature Publishing Group; 2005;439: 584. Available: https://doi.org/10.1038/nature04328

23. Public Health England. Detection of *Salmonella* species. Microbiology Services. Food, Water & Environmental Microbiology Standard Method FNES16 (F13). Version 2 [Internet]. Version 2. Methods Working Group, editor. London: Public Health England; 2014. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/330682/National_SOP_FNES16_F13_Detection_of_*Salmonella*_Species.pdf

24. Malorny B, Paccassoni E, Fach P, Bunge C, Martin A, Helmuth R. Diagnostic real-time PCR for detection of *Salmonella* in food. Appl Environ Microbiol. United States; 2004;70: 7046–7052. doi:10.1128/AEM.70.12.7046-7052.2004

25. Beltran P, Musser JM, Helmuth R, Farmer JJ 3rd, Frerichs WM, Wachsmuth IK, et al. Toward a population genetic analysis of *Salmonella*: genetic diversity and

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

652 relationships among strains of serotypes *S*. Choleraesuis, *S*. Derby, *S*. Dublin, *S*.
653 Enteritidis, *S*. Heidelberg, *S*. Infantis, *S*. Newport, and *S*. Typhimurium. Proc Natl
654 Acad Sci U S A. United States; 1988;85: 7753–7757.
655 26. Beltran P, Plock SA, Smith NH, Whittam TS, Old DC, Selander RK. Reference
656 collection of strains of the *Salmonella* Typhimurium complex from natural populations.
657 J Gen Microbiol. England; 1991;137: 601–606. doi:10.1099/00221287-137-3-601
658 27. Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, et al. The
659 complete genome sequence of *Escherichia coli* K-12. Science. United States;
660 1997;277: 1453–1462.
661 28. Yoshida CE, Kruczkiewicz P, Laing CR, Lingohr EJ, Gannon VPJ, Nash JHE, et al.
662 The *Salmonella In Silico* Typing Resource (SISTR): An Open Web-Accessible Tool for
663 Rapidly Typing and Subtyping Draft *Salmonella* Genome Assemblies. PLoS One.
664 United States; 2016;11: e0147101. doi:10.1371/journal.pone.0147101
665 29. Alikhan N-F, Zhou Z, Sergeant MJ, Achtman M. A genomic overview of the population
666 structure of *Salmonella*. PLoS Genet. United States; 2018;14: e1007261.
667 doi:10.1371/journal.pgen.1007261
668 30. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using
669 exact alignments. Genome Biol. England; 2014;15: R46. doi:10.1186/gb-2014-15-3-
670 r46
671 31. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence
672 data. Bioinformatics. England; 2014;30: 2114–2120.
673 doi:10.1093/bioinformatics/btu170
674 32. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes:
675 a new genome assembly algorithm and its applications to single-cell sequencing. J
676 Comput Biol. United States; 2012;19: 455–477. doi:10.1089/cmb.2012.0021
677 33. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for
678 genome assemblies. Bioinformatics. England; 2013;29: 1072–1075.
679 doi:10.1093/bioinformatics/btt086

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

680 34. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. England;
681 2014;30: 2068–2069. doi:10.1093/bioinformatics/btu153
682 35. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid
683 large-scale prokaryote pan genome analysis. Bioinformatics. England; 2015;31:
684 3691–3693. doi:10.1093/bioinformatics/btv421
685 36. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, et al. SNP-sites:
686 rapid efficient extraction of SNPs from multi-FASTA alignments. Microb genomics.
687 England; 2016;2: e000056. doi:10.1099/mgen.0.000056
688 37. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
689 large phylogenies. Bioinformatics. England; 2014;30: 1312–1313.
690 doi:10.1093/bioinformatics/btu033
691 38. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and
692 annotation of phylogenetic and other trees. Nucleic Acids Res. England; 2016;44:
693 W242-5. doi:10.1093/nar/gkw290
694 39. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data,
695 information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids
696 Res. England; 2014;42: D199-205. doi:10.1093/nar/gkt1076
697 40. Srikumar S, Kröger C, Hébrard M, Colgan A, Owen S V, Sivasankaran SK, et al.
698 RNA-seq Brings New Insights to the Intra-Macrophage Transcriptome of *Salmonella*
699 Typhimurium. PLOS Pathog. Public Library of Science; 2015;11: e1005262. Available:
700 https://doi.org/10.1371/journal.ppat.1005262
701 41. Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schaffer AA.
702 Database indexing for production MegaBLAST searches. Bioinformatics. England;
703 2008;24: 1757–1764. doi:10.1093/bioinformatics/btn322
704 42. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler
705 transform. Bioinformatics. England; 2009;25: 1754–1760.
706 doi:10.1093/bioinformatics/btp324
707 43. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Alignment/Map format and SAMtools. Bioinformatics. England; 2009;25: 2078–2079. doi:10.1093/bioinformatics/btp352

44. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nature biotechnology. United States; 2011. pp. 24–26. doi:10.1038/nbt.1754

45. Matuschek E, Brown DFJ, Kahlmeter G. Development of the EUCAST disk diffusion antimicrobial susceptibility testing method and its implementation in routine microbiology laboratories. Clin Microbiol Infect. England; 2014;20: O255-66. doi:10.1111/1469-0691.12373

46. Andrews JM, Howe RA. BSAC standardized disc susceptibility testing method (version 10). J Antimicrob Chemother. England; 2011;66: 2726–2757. doi:10.1093/jac/dkr359

47. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother. England; 2012;67: 2640–2644. doi:10.1093/jac/dks261

48. Magiorakos A-P, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, et al. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. Clin Microbiol Infect. England; 2012;18: 268–281. doi:10.1111/j.1469-0691.2011.03570.x

49. Hendriksen RS, Vieira AR, Karlsmose S, Lo Fo Wong DMA, Jensen AB, Wegener HC, et al. 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. United States; 2011;8: 887–900. doi:10.1089/fpd.2010.0787

50. Dallman T, Ashton P, Schafer U, Jironkin A, Painset A, Shaaban S, et al. SnapperDB: A database solution for routine sequencing analysis of bacterial isolates. Bioinformatics. England; 2018; doi:10.1093/bioinformatics/bty212

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

736 51. Fuller CC, Jawahir SL, Leano FT, Bidol SA, Signs K, Davis C, et al. A multi-state
737 *Salmonella* Typhimurium outbreak associated with frozen vacuum-packed rodents
738 used to feed snakes. Zoonoses Public Health. Germany; 2008;55: 481–487.
739 52. Kanagarajah S, Waldram A, Dolan G, Jenkins C, Ashton PM, Carrion Martin AI, et al.
740 Whole genome sequencing reveals an outbreak of *Salmonella* Enteritidis associated
741 with reptile feeder mice in the United Kingdom, 2012-2015. Food Microbiol. England;
742 2018;71: 32–38. doi:10.1016/j.fm.2017.04.005
743 53. Goldman TF. SmartDraw: A Hands-On Tutorial and Guide. 1st ed. Upper Saddle
744 River, NJ, USA: Prentice Hall Press; 2009.
745 54. Goupil BA, Trent AM, Bender, Jeff B, Olsen, Karen E, Morningstar BR,
746 Wuenschmann A. A Longitudinal Study of *Salmonella* from Snakes Used in a Public
747 Outreach Program. J Zoo Wildl Med. 2012;43: 836–841.
748 55. Feasey NA, Hadfield J, Keddy KH, Dallman TJ, Jacobs J, Deng X, et al. Distinct
749 *Salmonella* Enteritidis lineages associated with enterocolitis in high-income settings
750 and invasive disease in low-income settings. Nat Genet. United States; 2016;48:
751 1211–1217. doi:10.1038/ng.3644
752 56. Harrison RA. Snake Envenoming: A Disease of Poverty. 2009.
753 doi:10.1371/journal.pntd.0000569
754 57. Warrell DA, Greenwood BM, Davidson NM, Ormerod LD, Prentice CR. Necrosis,
755 haemorrhage and complement depletion following bites by the spitting cobra (Naja
756 nigricollis). Q J Med. England; 1976;45: 1–22.
757 58. Ogunfowokan O. Bite-to-hospital time and morbidity in victims of viper bite in a rural
758 hospital in Nigeria. African J Prim Heal Care Fam Med. AOSIS OpenJournals; 2012;4:
759 371. doi:10.4102/phcfm.v4i1.371
760 59. Chaudhuri RR, Morgan E, Peters SE, Pleasance SJ, Hudson DL, Davies HM, et al.
761 Comprehensive assignment of roles for *Salmonella* Typhimurium genes in intestinal
762 colonization of food-producing animals. PLoS Genet. United States; 2013;9:
763 e1003456. doi:10.1371/journal.pgen.1003456

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

764 60. Kostiala AA, Westerstrahle M, Muttilainen M. Neonatal *Salmonella* Panama infection
765 with meningitis. Acta Paediatr. Norway; 1992;81: 856–858. Available:
766 https://actavet.vfu.cz/83/4/0287/
767 61. Seif Y, Kavvas E, Lachance J-C, Yurkovich JT, Nuccio S-P, Fang X, et al. Genome-
768 scale metabolic reconstructions of multiple *Salmonella* strains reveal serovar-specific
769 metabolic traits. Nat Commun. 2018;9: 3771. doi:10.1038/s41467-018-06112-5
770 62. Nichols RJ, Sen S, Choo YJ, Beltrao P, Zietek M, Chaba R, et al. Phenotypic
771 landscape of a bacterial cell. Cell. United States; 2011;144: 143–156.
772 doi:10.1016/j.cell.2010.11.052
773 63. Gutnick D, Calvo JM, Klopotowski T, Ames BN. Compounds which serve as the sole
774 source of carbon or nitrogen for *Salmonella* Typhimurium LT-2. J Bacteriol. United
775 States; 1969;100: 215–219. Available:
776 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC315380/pdf/jbacter00586-0249.pdf
777 64. Kröger C, Fuchs TM. Characterization of the myo-inositol utilization island of
778 *Salmonella enterica* serovar Typhimurium. J Bacteriol. United States; 2009;191: 545–
779 554. doi:10.1128/JB.01253-08
780 65. Lamas A, Miranda JM, Regal P, Vázquez B, Franco CM, Cepeda A. A comprehensive
781 review of non-*enterica* subspecies of *Salmonella enterica*. Microbiol Res. 2018;206:
782 60–73. doi:https://doi.org/10.1016/j.micres.2017.09.010
783 66. McDonough PL, Shin SJ, Lein DH. Diagnostic and public health dilemma of lactose-
784 fermenting *Salmonella enterica* serotype Typhimurium in cattle in the Northeastern
785 United States. J Clin Microbiol. United States; 2000;38: 1221–1226. Available:
786 http://jcm.asm.org/content/38/3/1221.full.pdf+html
787 67. Parsons SK, Bull CM, Gordon DM. Spatial Variation and Survival of *Salmonella*
788 *enterica* Subspecies in a Population of Australian Sleepy Lizards (Tiliqua rugosa).
789 Appl Environ Microbiol. United States; 2015;81: 5804–5811. doi:10.1128/AEM.00997-
790 15
791 68. Nolle N, Felsl A, Heermann R, Fuchs TM. Genetic Characterization of the Galactitol

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

792 Utilization Pathway of *Salmonella enterica* Serovar *Typhimurium*. Becker A, editor.
793 Journal of Bacteriology. 1752 N St., N.W., Washington, DC; 2017.
794 doi:10.1128/JB.00595-16
795 69. Langridge GC, Fookes M, Connor TR, Feltwell T, Feasey N, Parsons BN, et al.
796 Patterns of genome evolution that have accompanied host adaptation in *Salmonella*.
797 Proc Natl Acad Sci U S A. United States; 2015;112: 863–868.
798 doi:10.1073/pnas.1416707112
799 70. Keebaugh AC, Thomas JW. The genomes of the South American opossum
800 (Monodelphis domestica) and platypus (Ornithorhynchus anatinus) encode a more
801 complete purine catabolic pathway than placental mammals. Comp Biochem Physiol
802 Part D Genomics Proteomics. Netherlands; 2009;4: 174–178.
803 doi:10.1016/j.cbd.2009.02.003
804 71. Gieraltowski L, Julian E, Pringle J, Macdonald K, Quilliam D, Marsden-Haug N, et al.
805 Nationwide outbreak of *Salmonella* Montevideo infections associated with
806 contaminated imported black and red pepper: warehouse membership cards provide
807 critical clues to identify the source. Epidemiol Infect. England; 2013;141: 1244–1252.
808 doi:10.1017/S0950268812001859
809 72. Gaffga NH, Barton Behravesh C, Ettestad PJ, Smelser CB, Rhorer AR, Cronquist AB,
810 et al. Outbreak of salmonellosis linked to live poultry from a mail-order hatchery. N
811 Engl J Med. United States; 2012;366: 2065–2073. doi:10.1056/NEJMoa1111818
812 73. Stocker P, Rosner B, Werber D, Kirchner M, Reinecke A, Wichmann-Schauer H, et al.
813 Outbreak of *Salmonella* Montevideo associated with a dietary food supplement
814 flagged in the Rapid Alert System for Food and Feed (RASFF) in Germany, 2010.
815 Euro Surveill Bull Eur sur les Mal Transm = Eur Commun Dis Bull. Sweden;
816 2011;16: 20040.
817 74. Kingsley RA, Msefula CL, Thomson NR, Kariuki S, Holt KE, Gordon MA, et al.
818 Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in
819 sub-Saharan Africa have a distinct genotype. Genome Res. United States; 2009;19:

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

2279–2287. doi:10.1101/gr.091017.109

75. Okoro CK, Barquist L, Connor TR, Harris SR, Clare S, Stevens MP, et al. Signatures of adaptation in human invasive *Salmonella* Typhimurium ST313 populations from sub-Saharan Africa. PLoS Negl Trop Dis. United States; 2015;9: e0003611. doi:10.1371/journal.pntd.0003611

Supporting Information Captions

S1 Table. Reptile-associated-*Salmonella* strain metadata and control strains metadata. The table includes all metadata associated with the venomous snake, non-venomous reptile and control strains used throughout this research including details on accessing raw reads were applicable.

S2 Table. Contextual reference genomes metadata and accession Numbers. The table lists the strains used in phylogenetic analysis to put the study strains into wider context and includes some basic information in addition to accession numbers.

S3 Table. QUAST assembly statistics. The table provides details on assembly statistics for each sequenced genome and highlights those strains which failed quality control in red.

S4 Table. Carbon source specific requirements. The table contains a list of the carbon sources used to determine the metabolic capacity of reptile associated *Salmonella.* Specific conditions, concentrations and references are given for each.

S1 Text. Carbon source utilisation gene sequences. A full list of the sequences used to determine the presence or absence of genes involved in carbon source utilisation in reptile associated *Salmonella.*

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

847 **S1 Figure. The diversity of *Salmonella* isolated from a collection of venomous snakes**
848 **and non-venomous reptiles including *S. bongori* outgroup.** Core genome maximum
849 likelihood phylogenetic tree. Tree was rooted using *S. bongori* shown here. 25 contextual
850 reference genomes representing previously sequenced isolates from each *Salmonella*
851 subgroup are indicated in purple.

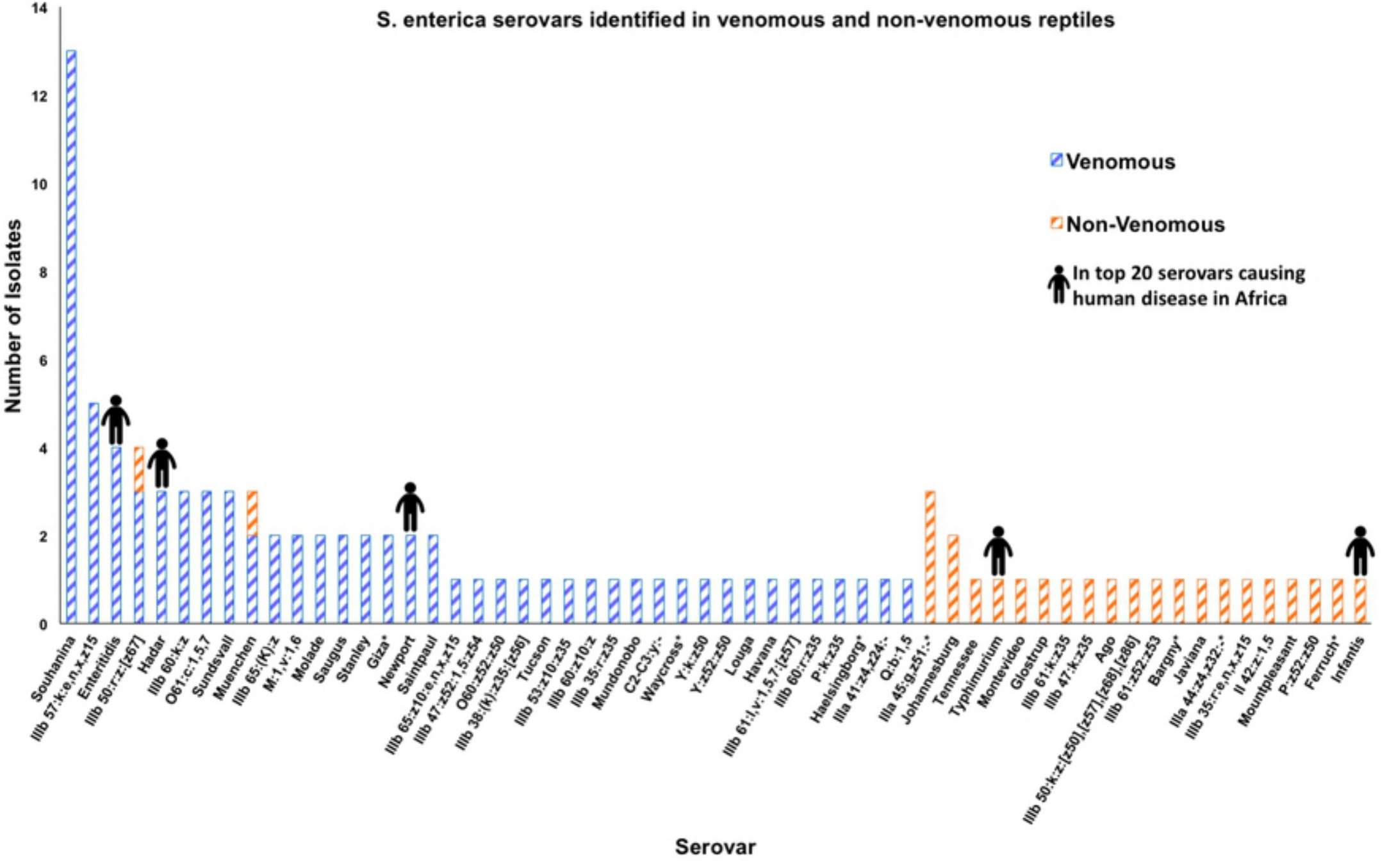

S. enterica serovars identified in venomous and non-venomous reptiles
Number of Isolates
0
2
4
6
8
10
12
14
Venomous
Non-Venomous
In top 20 serovars causing human disease in Africa
Serovar
Souhanina
IIIb 57:k:e,n,x,z15
Enteritidis
IIIb 50:r:z:[z67]
Hadar
IIIb 60:k:z
O61:c:1,5,7
Sundsvall
Muenchen
IIIb 65:(K):z
M:1,v:1,6
Molade
Saugus
Stanley
Giza*
Newport
Saintpaul
IIIb 65:z10:e,n,x,z15
IIIb 47:z52:1,5:z54
O60:z52:z50
IIIb 38:(k):z35:[z56]
Tucson
IIIb 53:z10:z35
IIIb 60:z10:z
IIIb 35:r:z35
Mundonobo
C2-C3:y:-
Waycross*
Y:k:z50
Y:z52:z50
Louga
Havana
IIIb 61:l,v:1,5,7:[z57]
IIIb 60:r:z35
P:k:z35
Haelsingborg*
IIIa 41:z4,z24:-
Q:b:1,5
IIIa 45:g,z51:-*
Johannesburg
Tennessee
Typhimurium
Montevideo
Glostrup
IIIb 61:k:z35
IIIb 47:k:z35
Ago
IIIb 50:k:z:[z50],[z57],[z68],[z86]
IIIb 61:z52:z53
Bargny*
Javiana
IIIa 44:z4,z32:-*
IIIb 35:r:e,n,x,z15
II 42:z:1,5
Mountpleasant
P:z52:z50
Ferruch*
Infantis

Figure 1

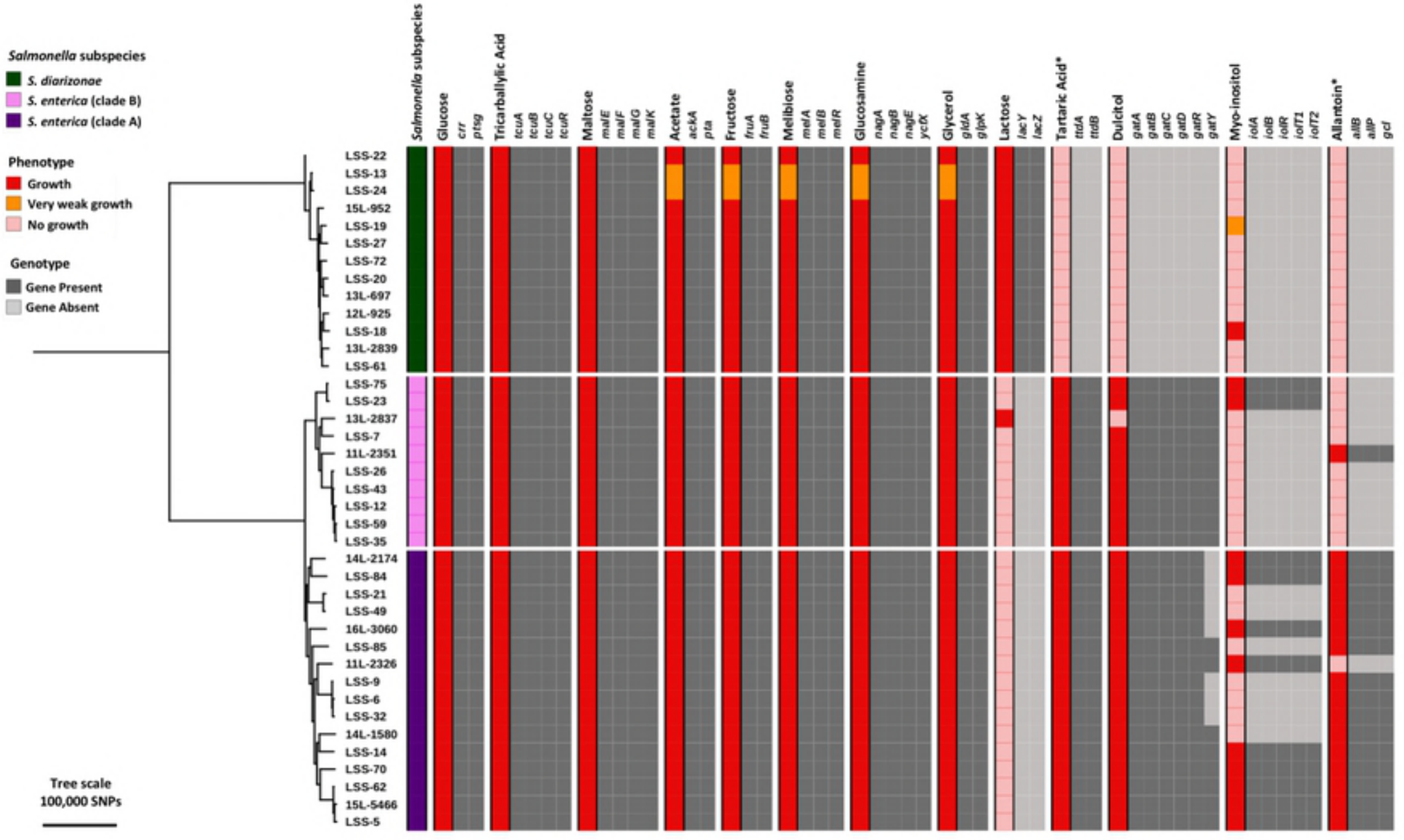

Salmonella subspecies
S. diarizonae
S. enterica (clade B)
S. enterica (clade A)
Phenotype
Growth
Very weak growth
No growth
Genotype
Gene Present
Gene Absent
Tree scale
100,000 SNPs
Salmonella subspecies
Glucose
crr
ptsg
Tricarballylic Acid
tcuA
tcuB
tcuC
tcuR
Maltose
malE
malF
malG
malK
Acetate
ackA
pta
Fructose
fruA
fruB
Melibiose
melA
melB
melR
Glucosamine
nagA
nagB
nagE
yciX
Glycerol
gldA
glpK
Lactose
lacY
lacZ
Tartaric Acid*
ttdA
ttdB
Dulcitol
gatA
gatB
gatC
gatD
gatR
gatY
Myo-inositol
iolA
iolB
iolR
iolT1
iolT2
Allantoin*
allB
allP
gcl
LSS-22
LSS-13
LSS-24
15L-952
LSS-19
LSS-27
LSS-72
LSS-20
13L-697
12L-925
LSS-18
13L-2839
LSS-61
LSS-75
LSS-23
13L-2837
LSS-7
11L-2351
LSS-26
LSS-43
LSS-12
LSS-59
LSS-35
14L-2174
LSS-84
LSS-21
LSS-49
16L-3060
LSS-85
11L-2326
LSS-9
LSS-6
LSS-32
14L-1580
LSS-14
LSS-70
LSS-62
15L-5466
LSS-5

Figure 3

bioRxiv preprint doi: https://doi.org/10.1101/526442; this version posted January 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

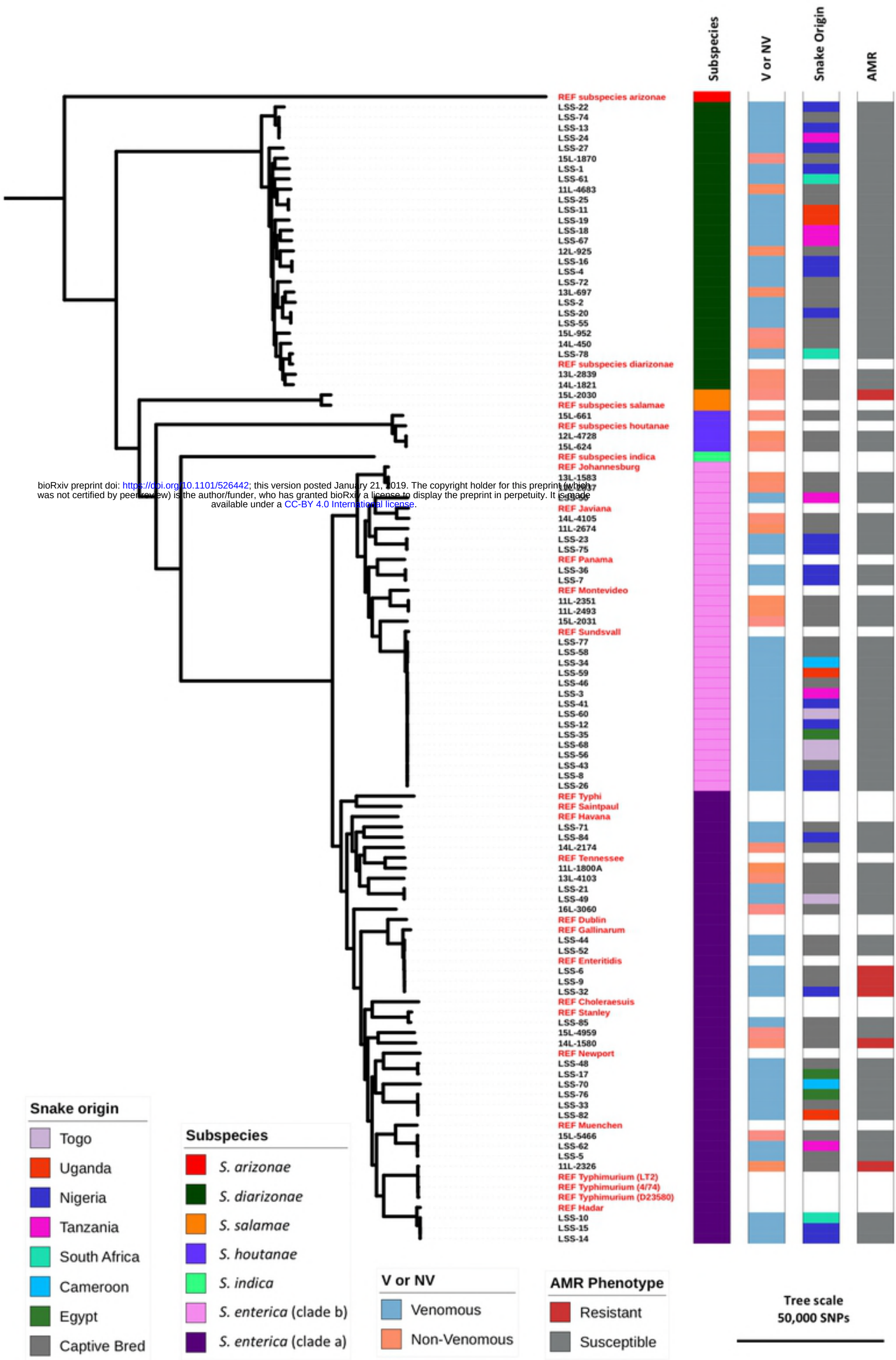


Figure 2