Draft Genome Sequences of 12 Monophasic *Salmonella enterica* subsp. *enterica* Serotype Typhimurium 1,4,[5], 12:i:– Strains Isolated from Wild Griffon Vultures in Eastern Spain

Clara Marin, a Giuseppe D’Auria, b,c Llúcia Martínez-Priego, b Francisco Marco-Jiménez a,d

a Instituto de Ciencias Biomédicas. Departamento de Producción Animal, Sanidad Animal y Ciencia y Tecnología de los Alimentos, Facultad de Veterinaria, Universidad Cardenal Herrera-CEU, CEU Universities, Alfara del Patriarca, Valencia, Spain
b Servicio de Secuenciación y Bioinformática, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO-Salud Pública), Valencia, Spain
c CIBER en Epidemiología y Salud Pública (CIBEResp), Madrid, Spain
d Instituto de Ciencia y Tecnología Animal, Universitat Politècnica de València, Valencia, Spain

ABSTRACT Monophasic *Salmonella enterica* subsp. *enterica* serovar Typhimurium is one of the most common zoonotic pathogens. *Salmonella* species reside in a wide variety of hosts, including wild animals. Thus, we report here the genome sequences of 12 monophasic *S.* Typhimurium strains isolated from healthy wild vultures to gain better insight into their epidemiology and host-pathogen interactions.

*Salmonella* spp. stand out as some of the most common causes of human bacterial food poisoning (1). Specifically, *Salmonella enterica* subsp. *enterica* serovar Typhimurium represented 21.8% of all reported serovars of confirmed human cases in 2016 in the European Union (1). Wild birds are vectors in the dissemination of livestock and human pathogens (2), including *Salmonella* spp. (3). Nevertheless, the number of wildlife species acting as reservoirs, amplifiers, and disseminators is unknown (4). *Salmonella* spp., including *S.* Typhimurium, have been isolated from vultures in several studies (2, 5, 6). Wild vulture monophasic *S.* Typhimurium strains displayed genomic DNA (gDNA) fingerprinting patterns similar to those observed in *Salmonella* strains from pig farms, suggesting that pig farms introduce *Salmonella* infection into vultures at supplementary feeding stations (5). Hence, genome analysis of multiple isolates of this *Salmonella* lineage should help to evaluate the potential risk to wildlife and the environment. To that end, we have sequenced the genomes of 12 isolates of monophasic *S.* Typhimurium strains isolated from wild vultures (Table 1).

**DNA extraction and sequencing.** Genomic DNA was isolated using a MagNA Pure LC DNA isolation kit III (Bacteria, Fungi; catalog number 03264785001). DNA libraries were generated using the Illumina Nextera XT library prep kit (catalog number FC-131-1024) and starting from 0.2 ng/µl of purified gDNA measured by a Qubit double-stranded DNA (dsDNA) high-sensitivity (HS) assay kit (catalog number Q32851). Sample multiplexing was performed using a Nextera XT index kit set C (catalog number FC-131-2003). Libraries were then sequenced using a 2 X 300-bp paired-end run (MiSeq reagent kit v3, catalog number MS-102-3001) on an Illumina MiSeq sequencer.

**Bioinformatics analysis.** Default parameters were used for all software, unless otherwise specified. The obtained raw sequencing data were quality checked using the PRINSEQ-lite program v0.20.4 (7), eliminating reads shorter than 50 bp, trim-
TABLE 1 Monophasic *Salmonella enterica* subsp. *enterica* serotype Typhimurium 1,4,[5],12:i− strains isolated from wild griffon vultures in eastern Spain

| Isolate          | Genome size (bp) | No. of reads | No. of contigs | Coverage (X) | \( N_{50} \) length (bp) (no. of contigs) | GC content (%) | ENA read accession no. | ENA assembly accession no. |
|------------------|------------------|-------------|---------------|-------------|------------------------------------------|---------------|------------------------|-----------------------------|
| Vulture-STm-CyP-1 | 5,162,468        | 806,314     | 67            | 63.47       | 461,225 (4)                             | 49            | ERR3385856             | SAMEA5530637                |
| Vulture-STm-CyP-2 | 5,265,851        | 669,207     | 84            | 51.01       | 442,846 (3)                             | 50            | ERR3385857             | SAMEA5530638                |
| Vulture-STm-CyP-3 | 5,336,318        | 862,984     | 66            | 69.78       | 423,460 (4)                             | 50            | ERR3385858             | SAMEA5530639                |
| Vulture-STm-CyP-4 | 5,026,433        | 1,175,547   | 58            | 91.76       | 461,244 (4)                             | 49            | ERR3385859             | SAMEA5530640                |
| Vulture-STm-CyP-5 | 4,559,258        | 1,270,926   | 57            | 101.19      | 492,138 (3)                             | 50            | ERR3385860             | SAMEA5530641                |
| Vulture-STm-CyP-6 | 5,344,058        | 945,925     | 75            | 74.92       | 416,712 (4)                             | 49            | ERR3385861             | SAMEA5530642                |
| Vulture-STm-CyP-7 | 5,026,635        | 909,679     | 56            | 71.52       | 423,380 (4)                             | 50            | ERR3385862             | SAMEA5530643                |
| Vulture-STm-CyP-8 | 5,076,141        | 666,708     | 62            | 50.92       | 423,460 (4)                             | 50            | ERR3385863             | SAMEA5530644                |
| Vulture-STm-CyP-9 | 5,025,060        | 798,412     | 56            | 59.62       | 461,235 (4)                             | 50            | ERR3385864             | SAMEA5530645                |
| Vulture-STm-CyP-10| 5,350,630        | 929,782     | 68            | 73.59       | 492,062 (4)                             | 50            | ERR3385865             | SAMEA5530646                |
| Vulture-STm-CyP-11| 5,025,905        | 835,164     | 57            | 64.13       | 461,239 (4)                             | 50            | ERR3385866             | SAMEA5530647                |
| Vulture-STm-CyP-12| 5,026,886        | 946,564     | 58            | 74.40       | 422,950 (4)                             | 50            | ERR3385867             | SAMEA5530648                |

The obtained reads were mapped against the *Salmonella enterica* subsp. *enterica* serovar Typhimurium SL1344 genome (GenBank accession number NC_016810) using the BWA-MEM algorithm (http://bio-bwa.sourceforge.net/). Unmapped reads were filtered out using SAMtools suite v1.8 (9) and assembled de novo using the MIRA program v3 (10). Contigs obtained from mapping against the reference genome have been extracted using BEDTools v2.24.0 (11), with the contig order following the mapping. De novo assembled contigs have been concatenated at the end of the assembly in order to collect complete genomic information. The final data sets were then annotated using Prokka v1.13 (12) to search for coding sequences (CDSs) annotated against the proper reference genome GenBank annotations, High-quality Automated and Manual Annotation of microbial Proteomes (HAMAP) v201701.18 (13), and Pfam v31.0 (14) databases. Table 1 summarizes the main genomic features.

The genomes ranged from 4.5 to 5.3 Mb in size, as described for other *Salmonella* strains (15). Sequencing generated an average G+ C content of 49.78%, which is similar to that reported previously for other *Salmonella* isolates (15). The number of contigs per assembly for each isolate ranged between 56 and 84. These genomes should allow for a detailed comparison of the attributes of monophasic *S. Typhimurium* isolates from wild vultures with those from other wild birds, livestock, and the environment, and with isolates contaminating the food chain.

**Data availability.** This whole-genome shotgun project has been deposited in the European Nucleotide Archive (ENA) under the project accession number PRJEB31693 and the sample accession numbers SAMEA5530637 to SAMEA5530648 (Table 1).

**ACKNOWLEDGMENTS**
Sequencing was carried out at the FISABIO Sequencing and Bioinformatics Service (Valencia, Spain).

This work was funded by Generalitat Valenciana (Government of València) and by CEU-UCH (Consolidación de Indicadores INDI15/16, INDI16/20, and INDI17/25).

**REFERENCES**
1. European Food Safety Authority (EFSA). 2017. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2016. EFSA J 15:5077.
2. Blanco G. 2018. Supplementary feeding as a source of multiresistant *Salmonella* in endangered Egyptian vultures. Transbound Emerg Dis 65:806. https://doi.org/10.1111/tbed.12806.
3. Krawiec M, Kuczkowski M, Kruzwicz A, Wieliczko A. 2015. Prevalence and genetic characteristics of *Salmonella* in free-living birds in Poland. BMC Vet Res 31:11–15. https://doi.org/10.1186/s12917-015-0332-x.
4. Molina-López RA, Vidal A, Obón E, Martín M, Darwich L. 2015. Multidrug-resistant *Salmonella enterica* serovar Typhimurium monophasic variant 4,12:i−− isolated from asymptomatic wildlife in a Catalonian wildlife rehabilitation center, Spain. J Wildl Dis 51:759–763. https://doi.org/10.7589/2015-01-019.
5. Marin C, Torres C, Marco-Jiménez F, Cerdà-Cuéllar M, Sevilla S, Ayats T, Vega S. 2018. Supplementary feeding stations for conservation of vultures could be an important source of monophasic *Salmonella Typhimurium* 1,4,[5],12:i−−. Sci Total Environ 636:449–455. https://doi.org/10.1016/j.scitotenv.2018.04.310.
6. Marin C, Palomeque MD, Marco-Jiménez F, Vega S. 2014. Wild griffon...
vultures (Gyps fulvus) as a source of Salmonella and Campylobacter in eastern Spain. PLoS One 9:e94191. https://doi.org/10.1371/journal.pone.0094191.
7. Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863–864. https://doi.org/10.1093/bioinformatics/btr026.
8. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
9. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map (SAM) format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
10. Chevreux B, Wetter T, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information, p 45–56. In Computer science and biology; proceedings of the German Conference on Bioinformatics, GCB ’99. GCB, Hannover, Germany.
11. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btp033.
12. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
13. Lima T, Auchincloss AH, Coudert E, Keller G, Michoud K, Rivoire C, Builliard V, de Castro E, Lachaize C, Baratin D, Phan I, Bougueleret L, Bairoch A. 2009. HAMAP: a database of completely sequenced microbial proteome sets and manually curated microbial protein families in UniProtKB/Swiss-Prot. Nucleic Acids Res 37:D471–D478. https://doi.org/10.1093/nar/gkn661.
14. Finn RD, Coggill P, Eberhardt RY, Eddy J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44:D279–D285. https://doi.org/10.1093/nar/gkv1344.
15. Seribelli AA, Frazão MR, Gonzales JC, Cao G, Leon MS, Kich JD, Allard MW, Falcão JP. 2018. Draft genome sequences of 20 Salmonella enterica subsp. enterica serovar Typhimurium strains isolated from swine in Santa Catarina, Brazil. Genome Announc 6:e00232-18. https://doi.org/10.1128/genomeA.00232-18.