# Population genetic structure of 4,12:a:- Salmonella enterica strains from harbor porpoises.

Jana K Haase, Derek J Brown, François-Xavier Weill, Henry Mather, Geoffrey Foster, Sylvain Brisse, John Wain, Mark Achtman

AEM
Journals.ASM.org

# Population Genetic Structure of 4,12:a:− *Salmonella enterica* Strains from Harbor Porpoises

**Jana K. Haase,[a] Derek J. Brown,[b] François-Xavier Weill,[c] Henry Mather,[b] Geoffrey Foster,[d] Sylvain Brisse,[c] John Wain,[e] and Mark Achtman[a]**

Environmental Research Institute, University College Cork, Cork, Ireland[a]; Scottish *Salmonella*, *Shigella* and *Clostridium difficile* Reference Laboratory, Microbiology Department, Stobhill Hospital, Glasgow, United Kingdom[b]; Institut Pasteur, Unité des Bactéries Pathogènes Entériques, Paris, France[c]; SAC Consulting Veterinary Services, Inverness, United Kingdom[d]; and Health Protection Agency, Colindale, London, United Kingdom[e]

**According to pulsed-field gel electrophoresis (PFGE) typing, 4,12:a:− *Salmonella enterica* isolates from harbor porpoises are highly diverse. However, porpoise isolates belong to only two multilocus sequence types within the eBurst group 18 (eBG18) genetic cluster, which also includes *S. enterica* serovars Bispebjerg and Abortusequi. Isolates of other, serologically similar serovars belong to unrelated eBGs. These assignments to eBGs were supported by eBG-specific sequences of the flagellar gene *fliC*.**

Isolates of 4,12:a:− *Salmonella enterica* subspecies *enterica* are frequently found in the pulmonary tract and other tissues of healthy and diseased harbor porpoises (*Phocoena phocoena*) from northern European coasts (4, 5, 8, 17). This monophasic antigenic formula corresponds to *S. enterica* serovar Fulica (6), which was first isolated from poultry in the United States in 1957 and only rarely thereafter, except for the harbor porpoise isolates (see Table S1 in the supplemental material). Twenty-seven percent of harbor porpoises in the United Kingdom are infected with 4,12:a:− strains (17), which are most commonly isolated from lung tissue (5, 17). It is possible that 4,12:a:− *Salmonella* isolates are transmitted by nematodes (3, 5), which are major porpoise parasites (7, 8, 13).

Little is known about the population structure of 4,12:a:− *Salmonella* from harbor porpoises (2, 10). Pulsed-field gel electrophoresis (PFGE) XbaI fingerprinting of 230 4,12:a:− isolates from harbor porpoises revealed 157 discrete fingerprint profiles (with a range of 10 to 20 fragments), of which 87% (136 profiles) were found only once (Fig. 1; see Table S2 in the supplemental material; for methods, see the supplemental material). Multiple isolates with uniform profiles were identified in three animals, and 3/157 profiles grouped strains from multiple animals. The latter profiles spanned periods of isolation of 1, 9, and 13 years. However, 15/20 porpoises with multiple isolates yielded two or more profiles. Furthermore, repeating PFGE typing of multiple single colonies from slopes, which were inoculated from a single colony and stored at room temperature for several years, revealed multiple PFGE profiles that had probably accumulated during laboratory storage (see Fig. S1 in the supplemental material).

Our PFGE observations raise the question of whether 4,12:a:− isolates from harbor porpoises are phylogenetically related or comprise multiple lineages. Previous multilocus sequence typing (MLST) analyses of >4,000 *S. enterica* subspecies *enterica* isolates have shown that numerous serovars were polyphyletic, comprising multiple unrelated sequence types (STs) or clusters of related STs (eBurst groups [eBGs]) (1, 12). MLST of 21 representatives of the PFGE diversity showed that all isolates belong to two STs (ST416, 20 isolates; ST417, 1 isolate) within eBG18 (Fig. 2; see Table S3 in the supplemental material). ST416 and ST417 differ by one single nucleotide polymorphism (SNP) in one of the seven MLST loci that were sequenced. Thus, 4,12:a:− *Salmonella* isolates from porpoises are genetically monomorphic according to MLST, similar to isolates of *S. enterica* serovar Typhi (9, 11). Comparisons with the MLST website revealed that eBG18 is genetically unrelated to other STs or eBGs, differing by at least five loci.

We also investigated the MLST patterns of 27 strains from seven other serovars with similar antigenic formulas (see Table S1 in the supplemental material). eBG18 also includes three other STs containing strains of *S. enterica* serovars Bispebjerg (1,4,[5],12:a:e.n.x), which has been isolated from turtles, humans, and birds, and Abortusequi (4,12:−:e,n,x), which causes disease in horses (Fig. 2; see Table S3 in the supplemental material). These serovars are found only in eBG18 according to the MLST website, and eBG18 contains no other serovars. Thus, eBG18 includes isolates with overlapping but discrete antigenic formulas.

Other serovars, including *S. enterica* serovar Fulica, which is antigenically identical to the porpoise isolates, were assigned to multiple, unrelated eBGs and STs (Fig. 2; see Table S3 in the supplemental material). *S. enterica* serovar Arechavaleta (4,[5],12:a:1,7) is polyphyletic, because seven isolates were distributed among four eBGs or STs. *S. enterica* serovar Hessarek (4,12,[27]:a:1,5) is also polyphyletic, with the two tested isolates belonging to two unrelated STs, one of which (ST255) is in the same eBG (eBG74) as the reference Fulica strain. Finally, single isolates of *S. enterica* serovars Kisangani and Tinda were assigned to two other STs. Thus, serological relatedness did not correlate with genetic relatedness for these serovars.

The antigenic formulas of all the serovars investigated here share similar lipopolysaccharide (LPS) epitopes and the flagel-

**Received** 23 July 2012 **Accepted** 24 September 2012
**Published ahead of print** 5 October 2012
Address correspondence to Mark Achtman, m.achtman@ucc.ie.
Supplemental material for this article may be found at http://aem.asm.org/.

doi:10.1128/AEM.02310-12

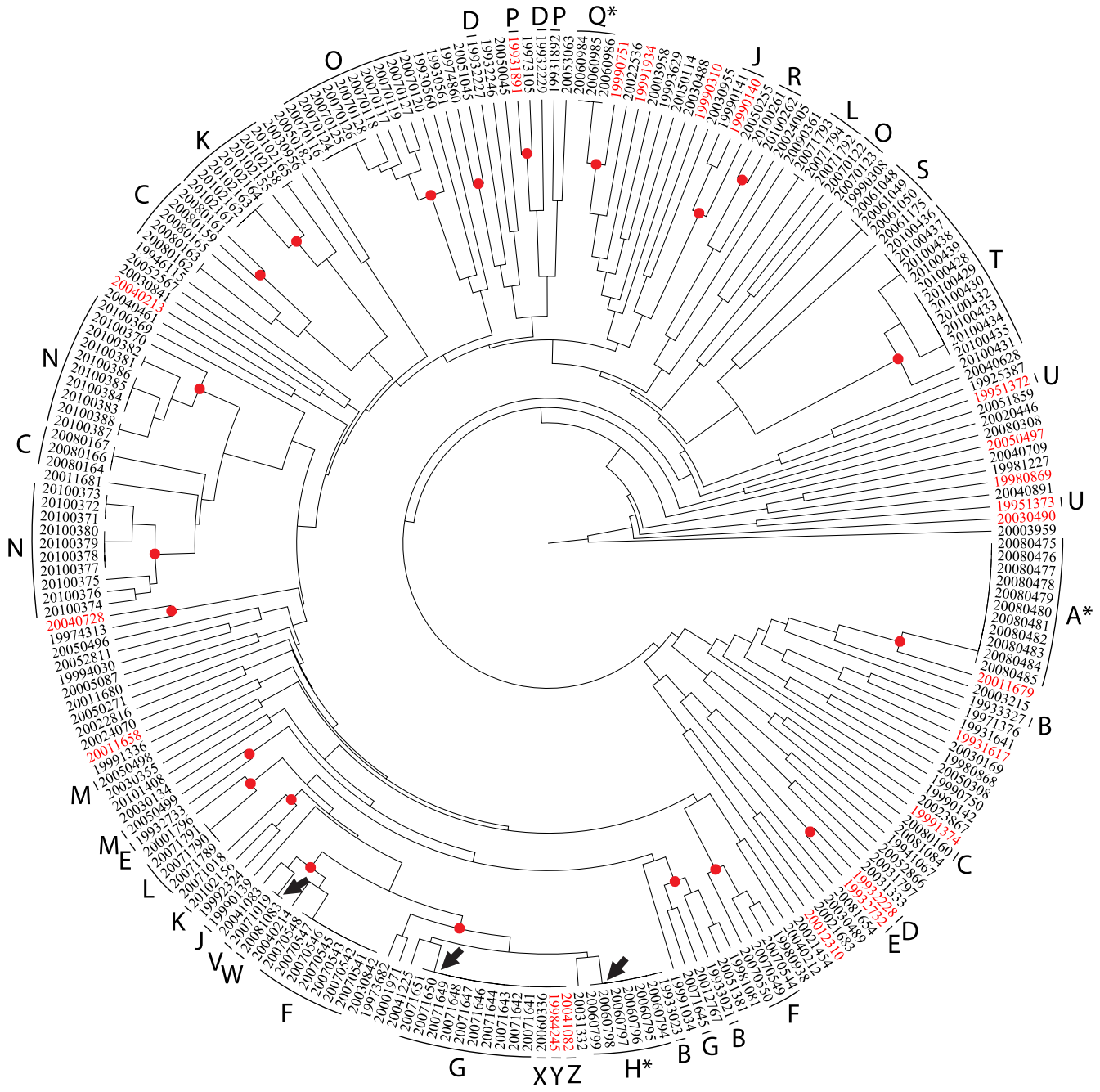



**FIG 1** Relationships of PFGE profiles for 230 *S. enterica* isolates from harbor porpoises according to an unweighted-pair group method using average linkages (UPGMA) dendrogram. The animal source is indicated for selected isolates by capital letters outside the dendrogram which designate the porpoise from which strains were isolated. Letters are followed by an asterisk for porpoises from which all isolated strains possessed identical PFGE profiles. Arrows indicate identical PFGE profiles in multiple porpoises. The strain designations of isolates that were subjected to MLST analysis are highlighted in red, whereas red dots indicate nodes connecting PFGE profiles that differed by fewer than four band differences.

lar phase 1 "a" epitope, except that *S. enterica* serovar Abortusequi does not express phase 1 (see Table S1 in the supplemental material). We sequenced 1,400 bp of *fliC* (phase 1) from all strains tested by MLST. *fliC* was identical in all *S. enterica* serovar Bispebjerg and 4,12:a:− harbor porpoise isolates (Fig. 3). Even though the FliC phase 1 antigen is not expressed in *S. enterica* serovar Abortusequi, *fliC* of 6/7 strains differed from that sequence by only one (nonsynonymous) SNP. We could not amplify *fliC* from the seventh strain (8259/94). *fliC* from all other isolates differed at multiple SNPs. Most of these SNPs were due to synonymous polymorphisms, resulting in only 10 polymorphic amino acids in the entire collection and amino acid sequences identical to that of *S. enterica* serovar Bispebjerg in four isolates of *S. enterica* serovars Hessarek, Arechavaleta, and Tinda. The ratio of nonsynonymous to synonymous rates of substitution, ω, was low (0.06) and similar to ω for MLST

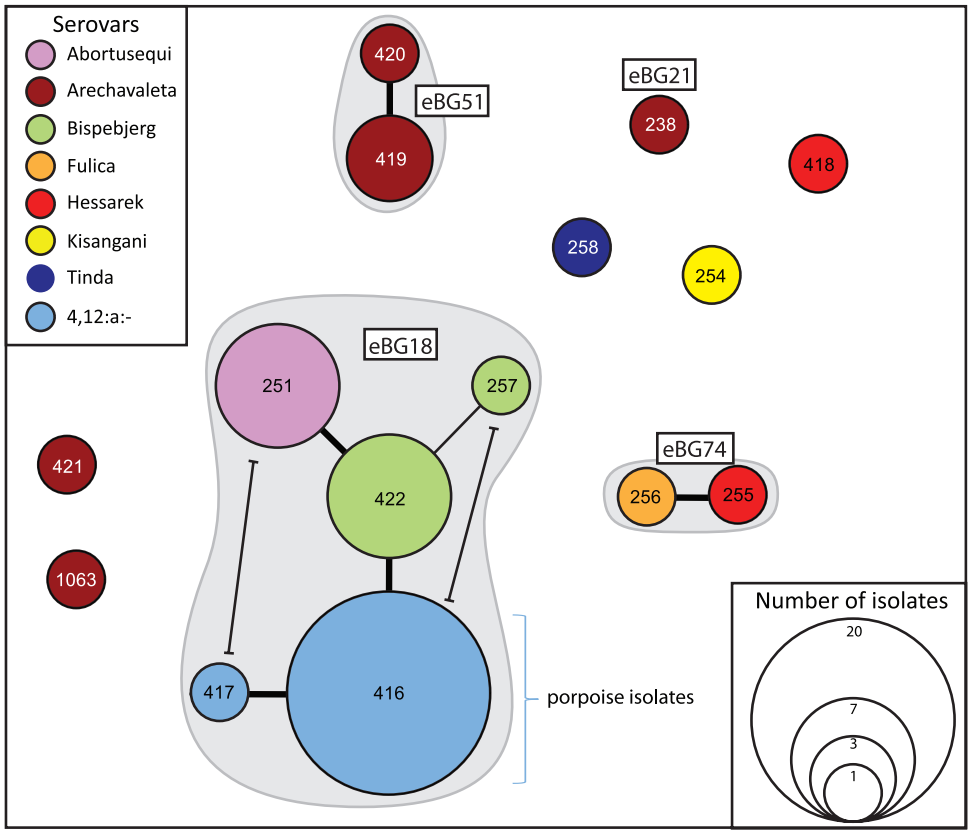


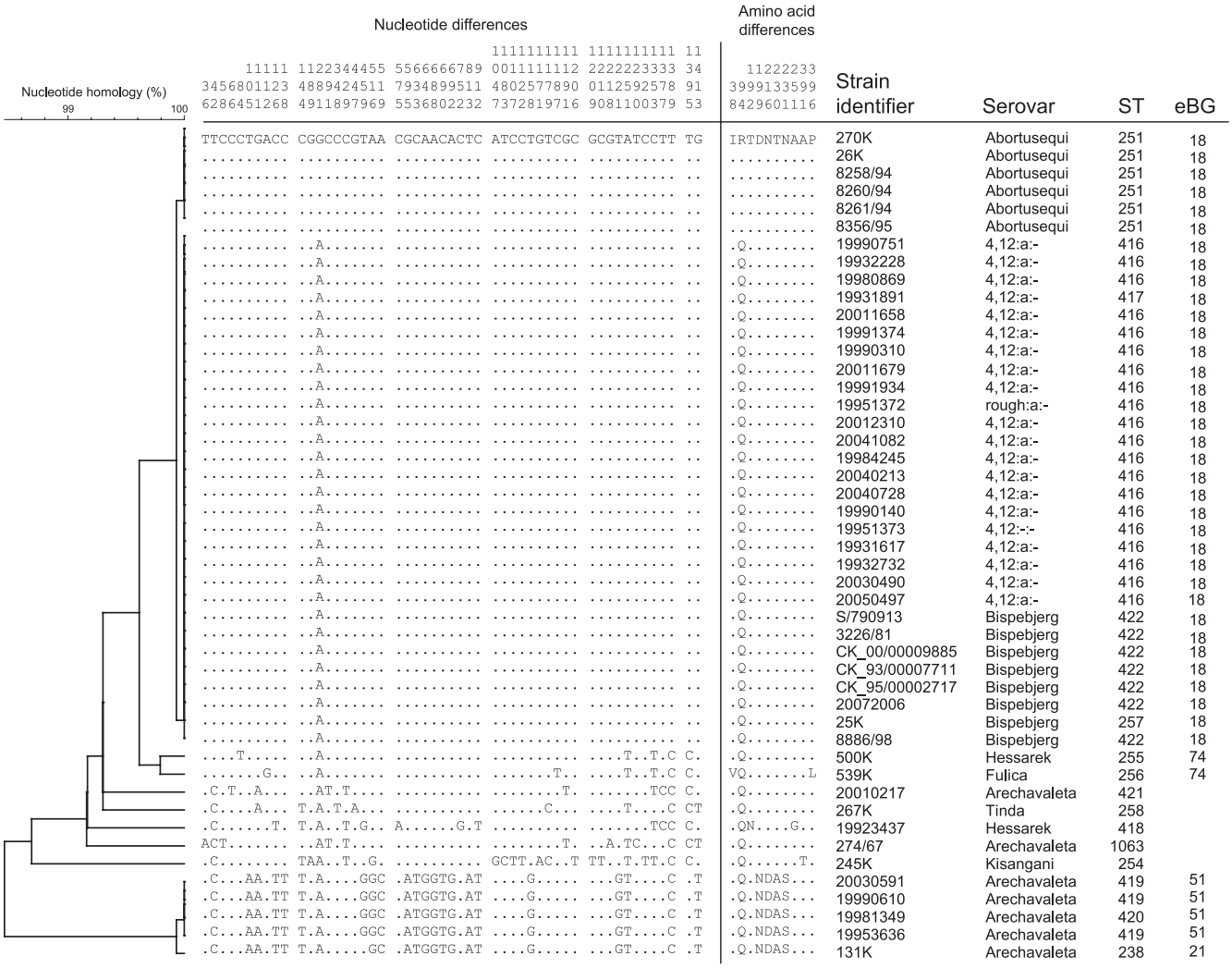

**FIG 2** Minimal spanning tree based on MLST alleles. Each ST is indicated by a circle encompassing the ST number, whose size reflects the number of isolates; eBGs are indicated by gray shading around STs, and the eBG designation is within a rectangle. STs are connected by thick black lines (six identical alleles) or thin black lines (five identical alleles). The primary connecting lines are not terminated, but two alternative connections at the level of five identical alleles are indicated by lines terminating in bars. All STs that are unconnected shared fewer than four alleles.

sequences (1), suggesting that nonsynonymous mutations are removed by purifying selection.

The primary differences between the serovars tested here are in the phase 2 antigen, which is 1,2, 1,5, 1,7, e,n,x, e,n,$z_{15}$, or not expressed (see Table S1 in the supplemental material). We sequenced 1,200 bp of *fljB* (phase 2) from multiple strains but could not amplify *fljB* from harbor porpoise isolates, with one exception (strain 20011679). The *fljB* amino acid sequences formed two discrete sequence clusters, encoding e,n,x or e,n,$z_{15}$ epitopes (*S. enterica* serovars Bispebjerg, Abortusequi, and Tinda) and 1,2, 1,5 or 1,7 epitopes (*S. enterica* serovars Hessarek, Arechavaleta, and Kisangani) (see Fig. S2 in the supplemental material). The *fljB* sequence from strain 20011679 differs by three amino acids and 14 or 15 SNPs from strains of *S. enterica* serovars Bispebjerg and Abortusequi (e,n,x), and 14 SNPs and one amino acid from an e,n,$z_{15}$ sequence from *S. enterica* serovar Tinda. BLASTN searches against *fljB* sequences encoding the e,n epitope in GenBank and unpublished sequences at the Institut Pasteur revealed that this 4,12:a:− sequence is identical to, or differs by only one SNP from, *fljB* of *S. enterica* serovars Brandenburg (4,[5],12:l,v:e,n,$z_{15}$) and Goettingen (9,12:l,v:e,n,$z_{15}$), both in eBG12 (see Fig. S3 in the supplemental material). Therefore, eBG12, which shares no alleles with eBG18, represents a potential source of the *fljB* in the exceptional 4,12:a:− isolate.

Our data show that the PFGE diversity reflects hypervariable changes that occur within individual animals and upon storage. According to MLST, 4,12:a:− *Salmonella* isolates from harbor porpoises form a genetically monomorphic group within eBG18 of *S. enterica* subspecies *enterica*, which also includes *S. enterica* serovars Bispebjerg and Abortusequi. In contrast, eBG18 isolates are unrelated to other *S. enterica* isolates with similar or identical serological patterns, including the reference strain for *S. enterica* serovar Fulica. We recommend that isolates from harbor porpoises be referred to by their ST or eBG designations rather than by their serological formulas, as the latter are not particularly informative. Our results also provide further support for the conclusion (1) that some eBGs are comprised of multiple serovars and some serovars conflate genetically unrelated bacteria. A particularly dramatic example of such conflation is *S. enterica* serovar Paratyphi B biovar Java (1), where distinct eBGs are differentially associated with infections of reptiles and poultry (15).

Members of eBG18 represent three distinct patterns of host adaptation and/or restriction (16), namely, infection of harbor porpoises, multiple hosts (*S. enterica* serovar Bispebjerg), and horses (*S. enterica* serovar Abortusequi). It would be interesting to perform MLST on an additional 4,12:a:− strain from a nematode parasitizing a harbor porpoise (3), because MLST distinguishes harbor porpoise isolates from antigenically identical *S. enterica* serovar Fulica. Similar patterns of host diversity have been reported within eBG4, which includes *S. enterica* serovar Enteritidis (1), which can infect a variety of hosts, as




| Strain identifier | Serovar | ST | eBG |
|---|---|---|---|
| 270K | Abortusequi | 251 | 18 |
| 26K | Abortusequi | 251 | 18 |
| 8258/94 | Abortusequi | 251 | 18 |
| 8260/94 | Abortusequi | 251 | 18 |
| 8261/94 | Abortusequi | 251 | 18 |
| 8356/95 | Abortusequi | 251 | 18 |
| 19990751 | 4,12:a:- | 416 | 18 |
| 19932228 | 4,12:a:- | 416 | 18 |
| 19980869 | 4,12:a:- | 416 | 18 |
| 19931891 | 4,12:a:- | 417 | 18 |
| 20011658 | 4,12:a:- | 416 | 18 |
| 19991374 | 4,12:a:- | 416 | 18 |
| 19990310 | 4,12:a:- | 416 | 18 |
| 20011679 | 4,12:a:- | 416 | 18 |
| 19991934 | 4,12:a:- | 416 | 18 |
| 19951372 | rough:a:- | 416 | 18 |
| 20012310 | 4,12:a:- | 416 | 18 |
| 20041082 | 4,12:a:- | 416 | 18 |
| 19984245 | 4,12:a:- | 416 | 18 |
| 20040213 | 4,12:a:- | 416 | 18 |
| 20040728 | 4,12:a:- | 416 | 18 |
| 19990140 | 4,12:a:- | 416 | 18 |
| 19951373 | 4,12:-:- | 416 | 18 |
| 19931617 | 4,12:a:- | 416 | 18 |
| 19932732 | 4,12:a:- | 416 | 18 |
| 20030490 | 4,12:a:- | 416 | 18 |
| 20050497 | 4,12:a:- | 416 | 18 |
| S/790913 | Bispebjerg | 422 | 18 |
| 3226/81 | Bispebjerg | 422 | 18 |
| CK_00/00009885 | Bispebjerg | 422 | 18 |
| CK_93/00007711 | Bispebjerg | 422 | 18 |
| CK_95/00002717 | Bispebjerg | 422 | 18 |
| 20072006 | Bispebjerg | 422 | 18 |
| 25K | Bispebjerg | 257 | 18 |
| 8886/98 | Bispebjerg | 422 | 18 |
| 500K | Hessarek | 255 | 74 |
| 539K | Fulica | 256 | 74 |
| 20010217 | Arechavaleta | 421 | |
| 267K | Tinda | 258 | |
| 19923437 | Hessarek | 418 | |
| 274/67 | Arechavaleta | 1063 | |
| 245K | Kisangani | 254 | |
| 20030591 | Arechavaleta | 419 | 51 |
| 19990610 | Arechavaleta | 419 | 51 |
| 19981349 | Arechavaleta | 420 | 51 |
| 19953636 | Arechavaleta | 419 | 51 |
| 131K | Arechavaleta | 238 | 21 |

**FIG 3** Nucleotide and amino acid diversity within *fliC* of serovars containing the "a" epitope. Left: unweighted-pair group method using average linkages (UPGMA) dendrogram of relatedness. An arbitrary reference nucleotide and amino acid sequence is indicated in the first line, and dots in other lines indicate identical nucleotides or amino acids. Nucleotide positions are given at the top. Right: information on each strain regarding serovar, ST, and eBG.

well as the host-adapted specialists *S. enterica* serovar Gallinarum and *S. enterica* serovar Gallinarum variant Pullorum, which cause two distinct avian diseases. A genomic comparison of one genome from each of the *S. enterica* serovars Enteritidis and Gallinarum indicated that specialization for avian disease was accompanied by genomic decay (14). Genomic analyses of eBG18 have not yet been performed but might be expected to yield interesting insights about host specificity in *Salmonella*.

## ACKNOWLEDGMENTS

We thank Christian Kornschober for supplying three strains and Sylvie Issenhuth-Jeanjean for her technical assistance.

The Scottish Strandings Scheme is operated under contract to the United Kingdom Department of Environment, Food and Rural Affairs (Defra), with financial support from the Scottish Government Marine Directorate. The project was funded by the Science Foundation Ireland (05/FE1/B882).

## REFERENCES

1. **Achtman M, et al.** 2012. Multilocus sequence typing as a replacement for serotyping in *Salmonella enterica.* PLoS Pathog. **8**:e1002776. doi:10.1371/journal.ppat.1002776.
2. **Crichton PB, Henry MS, Old DC.** 2000. Strain discrimination of a novel serotype of *Salmonella* from harbour porpoises (Phocoena phocoena) by molecular techniques. Vet. Microbiol. **76**:61–69.
3. **Davison N, Barnett J, Rule B, Chappell S, Wise G.** 2010. Group B *Salmonella* in lungworms from a harbour porpoise (*Phocoena phocoena*). Vet. Rec. **167**:351–352.
4. **Davison NJ, et al.** 2010. Prevalence of a host-adapted group B *Salmonella* enterica in harbour porpoises (*Phocoena phocoena*) from the south-west coast of England. Vet. Rec. **167**:173–176.
5. **Foster G, Patterson IA, Munro DS.** 1999. Monophasic group B *Salmonella* species infecting harbour porpoises (*Phocoena phocoena*) inhabiting Scottish coastal waters. Vet. Microbiol. **65**:227–231.
6. **Grimont PA, Weill F-X.** 2007. Antigenic formulae of the *Salmonella* serovars. WHO Collaborating Centre for Reference and Research on *Salmonella*, Paris, France.
7. **Jauniaux T, et al.** 2002. Post-mortem findings and causes of death of harbour porpoises (*Phocoena phocoena*) stranded from 1990 to 2000 along the coastlines of Belgium and Northern France. J. Comp. Pathol. **126**:243–253.
8. **Jepson PD, et al.** 2000. Pulmonary pathology of harbour porpoises (*Phocoena phocoena*) stranded in England and Wales between 1990 and 1996. Vet. Rec. **146**:721–728.
9. **Kidgell C, et al.** 2002. *Salmonella typhi*, the causative agent of typhoid fever, is approximately 50,000 years old. Infect. Genet. Evol. **2**:39–45.

10. **Old DC, Crichton PB, Taylor A, Mather H.** 2001. An attempt to identify the evolutionary origin of a novel serotype of *Salmonella enterica* isolated from harbour porpoises. J. Med. Microbiol. **50**:415–420.
11. **Roumagnac P, et al.** 2006. Evolutionary history of *Salmonella* Typhi. Science **314**:1301–1304.
12. **Sangal V, et al.** 2010. Evolution and population structure of *Salmonella enterica* serovar Newport. J. Bacteriol. **192**:6465–6476.
13. **Siebert U, et al.** 2001. Post-mortem findings in harbour porpoises (*Phocoena phocoena*) from the German North and Baltic Seas. J. Comp. Pathol. **124**:102–114.
14. **Thomson NR, et al.** 2008. Comparative genome analysis of *Salmonella* Enteritidis PT4 and *Salmonella* Gallinarum 287/91 provides insights into evolutionary and host adaptation pathways. Genome Res. **18**: 1624–1637.
15. **Toboldt A, et al.** 2012. Human infections attributable to the D-tartrate-fermenting variant of *Salmonella enterica* serovar Paratyphi B in Germany originate in reptiles and, on rare occasions, poultry. Appl. Environ. Microbiol. **78**:7347–7357.
16. **Uzzau S, et al.** 2000. Host adapted serotypes of *Salmonella enterica.* Epidemiol. Infect. **125**:229–255.
17. **Valderrama Vasquez CA, Macgregor SK, Rowcliffe JM, Jepson PD.** 2008. Occurrence of a monophasic strain of *Salmonella* group B isolated from cetaceans in England and Wales between 1990 and 2002. Environ. Microbiol. **10**:2462–2468.