Editor summary:

By analyzing the genome of over 9000 pig-associated isolates, this study shows that modernized agricultural systems have favored the acquisition of antimicrobial resistance genes, population expansion and global transmission of pig-enriched Salmonella over the past century.

1. Extended Data

Figure or Table #	Figure/Table title	Filename	Figure/Table Legend
Please group Extended Data items by type, in sequential order. Total number of items (Figs. + Tables) must not exceed 10.	One sentence only	Whole original file name including extension. i.e.: Smith_ED_Fig1.jpg	If you are citing a reference for the first time in these legends, please include all new references in the main text Methods References section, and carry on the numbering from the main References section of the paper. If your paper does not have a Methods section, include all new references at the end of the main Reference list.
Extended Data Fig. 1	**Circular presentation of the maximum-likelihood phylogeny in Figure 1A.**	Extended_Data_Fig_1 .eps	Outer rings: The clade, source, and geographic origin of each strain. Colored arcs underneath the tree show the cluster assignments as in the Key.
Extended Data Fig. 2	**Evaluation of the presence of temporal signal in *S. enterica* serovar Choleraesuis.**	Extended_Data_Fig_2 .eps	The analysis was performed on 587 genomes. (A) Linear regression between root-to-tip distances of strains and the sampling years with a coefficient of determination (R^2) of 0.67. (B) Substantially lower R^2 values were obtained for ten date-randomisation datasets. (C) The average (dots) and standard deviation (error bars) of the substitution rates for actual data

			(black) and ten date-randomization datasets (red), estimated by BactDating.
Extended Data Fig. 3	**Assessing the existence of temporal signal randomization test in pig-enriched ceBGs**	Extended_Data_Fig_3 .eps	(A) Coefficients of determination (R^2) for the ten date-randomisation tests were obtained by linear regression between root-to-tip distances of strains and the sampling years in pig-enriched ceBGs. (B) The average (dots) and standard deviations (error bars) of the substitution rates for actual data (black) and ten date-randomization datasets (red) by BactDating. Pig-enriched ceBGs including ceBG3 (n=3136), ceBG10 (n=622), ceBG17 (n=155), ceBG35 (n=516), ceBG37 (n=1012), ceBG40 (n=176), ceBG459 (n=1441), ceBG621 (n=787), ceBG1272 (n=586). The ancient sample was not included in ceBG1272 in both (A) and (B).
Extended Data Fig. 4	**The geographic states of the ancestral nodes in the Choleraesuis population after downsamplings.**	Extended_Data_Fig_4 .eps	(A) Downsampling tests of up to ten strains per country/region. (B) Downsampling tests of up to five strains per country/region. Both: the geographic states were predicted using TreeTime. Each test was run in 100 parallels. The pie charts illustrate the proportions of the best-supported geographic states in the 100 parallels for the corresponding ancestral nodes.
Extended Data Fig. 5	**Reconstructed ancestral host associations for all nodes in nine pig-enriched ceBGs**	Extended_Data_Fig_5 .eps	The corresponding ceBGs for each panel are: (A) ceBG3 (Derby), (B) ceBG1272 (Choleraesuis), (C) ceBG17 (Chailey), (D) ceBG35 (Worthington), (E) ceBG37 (London), (F) ceBG40 (Cerro), (G) ceBG459 (Johannesburg), (H) ceBG621 (Ohio), (I) ceBG10 (Adelaide). The mean values of effective population sizes with time were also shown for (A) and (B), with 95% confidence intervals in grey shapes. The orange and grey boxes show the periods of population expansion (A, B) or the periods of host transfers into pigs (C-I).
Extended Data Fig. 6	**The host transmission frequency of nine pig-enriched ceBGs.**	Extended_Data_Fig_6 .eps	The corresponding ceBGs for each panel are: (A) ceBG10 (Adelaide), (B) ceBG1272 (Choleraesuis), (C) ceBG17 (Chailey), (D) ceBG3 (Derby), (E) ceBG35 (Worthington), (F) ceBG37 (London), (G) ceBG 40 (Cerro), (H) ceBG 459 (Johannesburg),

			(I) ceBG 621 (Ohio). Different arrows represent the direction of host transmission, with darker colors indicating higher transfer frequency. "*" marks the most contributing host sources for the transmissions.
Extended Data Fig. 7	**A 5 X 5 table showing the summarised international transmission events of all nine pig-enriched ceBGs in *Salmonella* in the past 50 years.**	Extended_Data_Fig_7.eps	The maximum-likelihood phylogeny was reconstructed based on SNPs in the core genome of each ceBG and the date of origin was estimated using BactDating. TreeTime was applied to reconstruct the country sources for all nodes in the tree. A transmission was recorded when the ancestral node and descending node of a branch were different. All transmissions were then summarized and grouped based on their associated continents.
Extended Data Fig. 8	**Pearson's correlations analysis of sub-classification for pig-related products.**	Extended_Data_Fig_8.eps	The Sub classifications involved include (A) Pork offal, frozen (021011); (B) Pork, frozen (020649); (C) Pork, frozen cut (020329); (D) Pork, preserved (020322); (E) Pork offal, fresh (020630); (F) Pig fat (020910); (G) Pork, fresh cuted (020912); (H) Pork, fresh (020319); (I) Pig carcasses (020310); (J) Pig, live breeding (010391); (K) Pig, live, less 50kg (010392); (L) Pig, live, over 50kg (020311). The codes in parentheses represent the Harmonized System Codes of the products. Analyzing the correlation between the intercontinental dissemination of each pig-related commodity and the intercontinental transmission of pig-enriched ceBGs. R: Pearson's correlation coefficient.

14

1. Supplementary Information:

A. PDF Files

Item	Present?	Filename Whole original file name including extension. i.e.: Smith_SI.pdf. The extension must be .pdf	A brief, numerical description of file contents. i.e.: Supplementary Figures 1-4, Supplementary Discussion, and Supplementary Tables 1-4.

| Supplementary Information | Yes | Supplementary_Information.pdf | Supplementary Tables 1-4. |
| Reporting Summary | Yes | 6823_1_attach_25_30962.pdf | |

17

18

B. Additional Supplementary Files

Type	Number Each type of file (Table, Video, etc.) should be numbered from 1 onwards. Multiple files of the same type should be listed in sequence, i.e.: Supplementary Video 1, Supplementary Video 2, etc.	Filename Whole original file name including extension. i.e.: *Smith_Supplementary_Video_1.mov*	Legend or Descriptive Caption Describe the contents of the file
Supplementary Table	Additional Supplementary Files Table 1	Additional_Supplementary_Files.xlsx	In the manuscript and Additional Supplementary Files Table 1
Supplementary Table	Additional Supplementary Files Table 2	Additional_Supplementary_Files.xlsx	In the manuscript and Additional Supplementary Files Table 2

20

21

3. Source Data

Parent Figure or Table	Filename Whole original file name including extension. i.e.: *Smith_SourceData_Fig1.xls,* or *Smith_Unmodified_Gels_Fig1.pdf*	Data description i.e.: Unprocessed western Blots and/or gels, Statistical Source Data, etc.
Source Data Fig. 1	SourceData_Figure_1.xlsx	Statistical Source Data
Source Data Fig. 2	SourceData_Figure_2.xlsx; SourceData_Figure_2D_and_2F.txt; SourceData_Figure_2F.txt	Statistical Source Data
Source Data Fig. 3	SourceData_Figure_3A_and_ExtendedData_Figure_5.zip; SourceData_Figure3_and_ExtendedData_Figure6.xlsx	Statistical Source Data
Source Data Fig. 4	SourceData_Figure4BC_and_ExtendedData_Fig7.xlsx; SourceData_Figure4A_and_4DEFG_ExtendedData_Figure_8.xlsx	Statistical Source Data
Source Data Extended Data Fig./Table 1	SourceData_ExtendedData_Figure1.txt	Statistical Source Data
Source Data Extended Data Fig./Table 2	SourceData_ExtendedData_Figure2_and_Figure4.xlsx	Statistical Source Data

Source Data Extended Data Fig./Table 3	SourceData_ExtendedData_Figure3.xlsx	Statistical Source Data
Source Data Extended Data Fig./Table 4	SourceData_ExtendedData_Figure2_and_Figure4.xlsx	Statistical Source Data
Source Data Extended Data Fig./Table 5	SourceData_Figure3A_and_ExtendedData_Figure5.zip	Statistical Source Data
Source Data Extended Data Fig./Table 6	SourceData_Figure3_and_ExtendedData_Figure6.xlsx	Statistical Source Data
Source Data Extended Data Fig./Table 7	SourceData_Figure4BC_and_ExtendedData_Figure7.xlsx	Statistical Source Data
Source Data Extended Data Fig./Table 8	SourceData_Figure4A_4DEFG_and_ExtendedData_Figure8.xlsx	Statistical Source Data

23

24

Centralized industrialization of pork in Europe and America contributes to the global spread of *Salmonella enterica*

Heng Li [1,2], Yilei Wu [1,3], Dan Feng [1,2], Quangui Jiang [1,2], Shengkai Li [1], Jie Rong [1], Ling Zhong [1], Ulrich Methner [4], Laura Baxter [5], Sascha Ott [6], Daniel Falush[7], Zhenpeng Li[8], Xiangyu Deng[9], Xin Lu [8], Yi Ren [10], Biao Kan [8], Zhemin Zhou [1,2,8*]

Affiliations

[1] Key Laboratory of Alkene-carbon Fibres-based Technology & Application for Detection of Major Infectious Diseases, MOE Key Laboratory of Geriatric Diseases and Immunology, Pasteurien College, Suzhou Medical College, Soochow University, Suzhou, China

[2] Suzhou Key Laboratory of Pathogen Bioscience and Anti-infective Medicine, Jiangsu Province Engineering Research Center of Precision Diagnostics and Therapeutics Development, Soochow University, Suzhou, China

[3] Department of Biological Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, China

[4] Institute of Bacterial Infections and Zoonoses, Friedrich-Loeffler-Institut, Jena, Germany

[5] Warwick Bioinformatics Research Technology Platform, University of Warwick, Coventry, UK

[6] Warwick Medical School, University of Warwick, Coventry, UK

[7] The Center for Microbes, Development and Health, CAS Key Laboratory of Molecular Virology and Immunology, Shanghai Institute of Immunity and Infection, Chinese Academy of Sciences, Shanghai, China

[8] National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China

[9] Center for Food Safety, University of Georgia, Griffin GA, USA

[10] Iotabiome Biotechnology Inc., Suzhou, China

51   **Abstract**

52   *Salmonella enterica* (*S. enterica*) causes severe foodborne infections through contamination of

53   the food supply chain. Its evolution has been associated with human activities, especially animal

54   husbandry. Advances in intensive farming and global transportation have substantially reshaped

55   the pig industry, but their impact on the evolution of associated zoonotic pathogens such as *S.*

56   *enterica* remains unresolved. Here we investigated the population fluctuation, accumulation of

57   antimicrobial-resistant genes, and international serovar Choleraesuis transmission of nine pig-

58   enriched *S. enterica* populations comprising more than 9000 genomes. Most changes were found

59   to be attributable to the developments of the modern pig industry. All pig-enriched salmonellae

60   experienced host transfers in pigs and/or population expansions over the past century, with pigs

61   and pork having become the main sources of *S. enterica* transmissions to other hosts. Overall,

62   our analysis revealed strong associations between the transmission of pig-enriched salmonellae

63   and the global pork trade.

64

65

Main

Salmonella enterica infiltrates food supply chains through the contamination of food, water, or food-processing facilities[1], resulting in life-threatening foodborne infections with 108.1 million illnesses and 291,000 deaths annually[2]. Pork and pigs are prominent sources of *S. enterica* infections, accounting for ~31.1% of salmonellosis and 9.3% of disease outbreaks in the European Union[3]. Despite its acknowledged role in mediating transmissions and outbreaks of viral diseases[4,5], the contribution of pigs to the global dissemination of bacterial pathogens, including *S. enterica*, remains insufficiently explored within the framework of the "One Health" strategy.

The developments of intensive farming and global trade over the past century have drastically transformed pig agriculture[6], giving rise to two industrial hubs, Europe and the US, that collectively represent >32% of international pork and pig breed trades[7,8]. While most of the >2000 populations in *S. enterica*, recognized by their serovars, eBGs, or ceBGs based on the genomic sequences[9], are ubiquitous, some populations predominantly comprise strains associated with pigs[10] and are found to spread regionally through movements of pigs and/or wild boars[11]. Nonetheless, it remains unclear how pathogens, especially these pig-enriched salmonellae, have disseminated globally and how their population dynamics have been recrafted by modern agriculture.

Results

Landscape of pig-enriched *Salmonella enterica* populations

Systematic investigation of all 362,931 *Salmonella* strains publicly accessible in EnteroBase (July 2022) showed that pigs/pork accounts for 17,623 strains (4.9%) in 252 ceBGs and is the 2nd most frequent livestock source of *S. enterica* after poultry over the past century (Fig. 1A, B). There are 61 major ceBGs that each contains ≥20 pig-related strains, of which nine have ≥40% of their strains from pigs/pork (Fig. 1C, D; Additional Supplementary Files Table 1), including prominent pig-enriched populations such as ceBG1272 (Choleraesuis) and ceBG3 (Derby), and others like ceBG10 (Adelaide) and ceBG459 (Johannesburg). Pig is the primary source (4393; 49%) of strains in these 9 pig-enriched ceBGs, followed by humans (2482; 28%) and other

94   animals (1279; 14%). The other ceBGs have lower levels of pig- associated strains (0.2-38%)

95   and exhibit no clear host preference, including ceBG2 (Typhimurium) which has only 6% (4612)

96   of pig-associated strains (Fig. 1C, D).

97   The HC5 clusters in EnteroBase, namely clusters of strains with ≤5 allelic differences in their

98   core genes, have been extensively used in epidemiological investigations for designating

99   genetically almost identical bacteria such as those from disease outbreaks[12]. Unexpectedly, while

100   most HC5s in the pig-enriched ceBGs are from single countries, there are 35 HC5s each

101   consisting of strains from ≥2 countries, including 15 HC5s with strains from different continents

102   (Supplementary Table S1), indicating very recent international or even cross-continental

103   transmissions. Notably, all these international HC5s contain at least one pig strain, prompting the

104   importance of pigs for these long-range transmissions.

105

106   **Europe as the main genetic repository of serovar Choleraesuis**

107   We reconstructed a maximum-likelihood phylogeny of serovar Choleraesuis based on 21,948 non-

108   repetitive, non-recombinant single nucleotide polymorphisms (SNPs) in the core genome and used it

109   to divide strains into three lineages of CS1, CS2, and CS3 from the root that were separated by 2932

110   to 12,435 SNPs (Additional Supplementary Files Table 2). Except for CS3 which contained only two

111   strains, the other two lineages were subdivided into clades and clusters (Fig. 2A and Extended Data

112   Fig. 1). CS1 consists of 8 clusters in two clades of CS1.1 and CS1.2, and CS2 consists of 19 clusters

113   in three clades of CS2.1 to CS2.3. High geographic and host specificities were found in certain clades

114   and clusters. For example, most of the strains in Chinese mainland and Vietnam fell into Clade 1.2,

115   while many of the British strains were from Clade 1.1. In addition, 92% of the US strains grouped

116   with those from Chinese Taiwan in Clade 2.2, while >60% of European wild boar strains were from

117   Clades 2.3.

118   We evaluated the genetic diversity of Choleraesuis in different regions worldwide. The Simpson's

119   diversity index (SDI) for the presence of clades in different regions (Fig. 2E) showed that West

120   and North European regions had the greatest levels of diversities (0.66 – 0.7), followed by South

121   America and East Asia (0.63). In contrast, Africa and North America had the lowest SDIs (0.14

122 and 0.17, respectively). Furthermore, to minimize the impact of oversampling in developed

123 countries, we performed a country-level comparison, which also showed that the countries from

124 North Europe exhibited greater levels of SDIs than those from other regions (Supplementary Table

125 S2). This suggested Europe as the main genetic repository and probably the origin of serovar

126 Choleraesuis (Fig. 2E).

127

Host-specific ARGs accumulation in Choleraesuis

129 Choleraesuis strains carried many more antimicrobial resistant genes (ARGs) than its human-

130 specific analog, Paratyphi C (Fig. 2B, C), which was separated from serovar Choleraesuis only

131 ~4000yrs ago, suggesting an association between hosts and ARGs. Furthermore, significant

132 differences in ARG levels were found between Choleraesuis strains from different hosts and

133 countries/regions. In particular, strains from humans and livestock carried ~9-fold more ARGs

134 than those from wild boars (Fig. 2C), and strains from Chinese Taiwan and Vietnam carried more

135 ARGs than those from others (Fig. 2D).

136 Moreover, Choleraesuis isolated from pigs exhibited high resistance against aminoglycoside,

137 sulfonamide, tetracycline, and beta-lactam, which were all common feed supplements for

138 intensive farming[12] (Fig. 2A). In contrast, resistances against clinical antimicrobials including

139 quinolone, trimethoprim, and cephalosporins were much fewer, and often found only in human

140 strains. Notably, the colistin-resistant genes, *mcr*-1 and *mcr*-3, commonly detected in pigs[13], were

141 found in eight human/pig strains from the UK, China, Brazil, and Germany, underscoring their

142 global presence (Supplementary Table S3). Additionally, continuous increases in ARG carriages

143 over time were spotted in CS1 and CS2 strains isolated after the 1970s (Fig. 2B), except that the

144 ARG carriages in CS2 dropped after the 2010s due to an increase of strains from wild boars.

145

International transmission of serovar Choleraesuis

147 Significant temporal signals were detected in serovar Choleraesuis, with and without the ancient

148 genotype (Extended Data Figs. 2 and 3). Bayesian inferences predicted that the most recent

149 common ancestor (MRCA) of Choleraesuis had probably been circulating in Europe before 2394

150 BP (CI95% 2276-2521 BP) and diverged there into CS1 and CS2 in 1785 and 1870, respectively

151 (Fig. 2F). The first predicted transmission outside of Europe occurred before 1893 (CI95% 1891-

152 1896) and resulted in Clade CS2.2 in the US. Soon after, the effective population size of

153 Choleraesuis was predicted to experience an expansion in the early 20th, coincident with the rapid

154 development of intensive pig farming, and reached its first peak in the 1930s, before the

155 commercial use of synthetic sulphonamides and other antimicrobials in animals[14].

156 After 20 years of stale, a second expansion of Choleraesuis was predicted between the 1950s and

157 the 1980s (Fig. 2G). The frequencies of international transmissions also increased, possibly

158 associated with the rapid expansion of global agricultural trade as part of the post-war waves of

159 livestock revolution and trade globalization[15]. Europe and the US were the major sources of

160 international transmissions. For example, European CS1 strains were repetitively transmitted to

161 Chinese mainland and Southeast Asia (Fig. 2F). Furthermore, CS2.2 strains were transmitted from

162 the US to Chinese Taiwan (Fig. 2F, CS2.2.4) in 1962 (CI95% 1957-1968) and became endemic

163 there for more than50 years, causing major human outbreaks between 1996 and 2002[16]. The

164 population size of Choleraesuis reached its peak in 1985 and underwent continuous decreases

165 afterward (Fig. 2G). This was also accompanied by a decrease in long-range transmissions,

166 although the local transmissions in Europe remained frequent, partially attributed to the movement

167 of wild boars[17].

168 The majority of the Choleraesuis genomes were isolated from North European countries, which

169 could lead to sampling bias in the analyses. Therefore, we performed phylogeographic

170 reconstructions by downsampling at most ten random genomes from each country (Extended Data

171 Fig. 4 and Supplementary Table S4). Summarising 100 downsampling results together, we found

172 that the MRCA of Choleraesuis, as well as the MRCAs of all major clades except for Clade 2.1,

173 were still from North Europe (Extended Data Fig. 4A). Downsampling to at most five genomes

174 per country still proposed North Europe as the origin of the whole population but made Peru the

175 origin of Lineage 1 (Extended Data Fig. 4B). These differences likely resulted from the fact that

176 there was only one cluster of strains for each of Peru and Cameroon near the basal of Clade 2.1

177 and Lineage 1, respectively.

178

Intensive farming in establishments of pig-enriched serovars

We further evaluated the role of pigs in the evolution of all nine pig-enriched populations. To this end, we demonstrated the presence of temporal signals in all nine populations by date randomisations (Extended Data Fig. 3) and estimated their temporal phylogenies and ancestral host transfers. Notably, apart from ceBG40 (Cerro) which originated around 1954, the MRCAs for other populations were all predicted to be present before the 19th century (Fig. 3A and Extended Data Fig. 5). However, except for ceBG3 (Derby) and ceBG1272 (Choleraesuis), other populations were originally present in hosts other than pigs, and only transferred into pigs after 1930. Furthermore, we evidenced at least eight human-to-pig transmissions in ceBG3 (Derby) during 1906-1942, resulting in the establishment of the contemporary pig-enriched lineages and major population expansion (Extended Data Fig. 5A). Thus, all pig-associated salmonellae, apart from Choleraesuis, likely experienced host transfers in the 20th century (Fig. 3A). Furthermore, the accumulation of pseudogenes has been associated with a drastic change, such as host adaptation of the bacteria. However, we did not evident an accumulation of pseudogenes in any population (Fig. 3D), except for the pig-adapted Choleraesuis which has >17.3% of its genes disrupted[7].

Different from the long-term trend of human-to-pig transfers, the majority (30-60%) of the recent host transfers in these populations, including those into humans, have been contributed by pigs in the past 50 years (Fig. 3B and Extended Data Fig. 6). Conversely, pigs contribute much less to host transfers in eight populations that had lower proportions (7-38%) of pig strains (Fig. 3C), demonstrating the importance of pigs from intensive farming as a hub of host transfers in the pig-enriched populations.

200

Dispersal of *Salmonella enterica* from Europe and America

The reconstructed international transmissions in all nine populations showed that 68-96% of transmissions into each continent were from either Europe or America (Fig. 4A), exhibiting similar patterns to the trade data of the pork-related products in the Harvard database (Fig. 4B). We then summarized the cross-continental trades for each of the 5,014 product categories in the

206 Harvard database. The transmission of pig-enriched salmonellae was shown in a 5×5 table, in
207 which rows and columns represented the continental sources and targets, and each cell showed the
208 percentage contribution of a source continent to the influx of the target (Extended Data Fig. 7).
209 This resulted in a dataset for trade/transmissions, of which the pairwise Pearson's correlations
210 were calculated and projected to a 2-D space using an unsupervised method, the uniform manifold
211 approximation and projection (UMAP)[18] (Fig. 4C).

212 The trade data for most of the animal-related products fell into two clusters in the projection. The
213 first cluster consisted of almost all bovine and poultry products as well as live pigs, and the second
214 cluster consisted of five pig-related products and one poultry product. Impressively, the
215 transmission of the pig-enriched salmonellae also fell in the second cluster, exhibiting 0.87-0.96
216 correlation coefficients to the pig-related products (p<0.0001) (Fig. 4C). Detailed investigation of
217 the products in the 2nd cluster suggested that they were either frozen or processed pork or offal and
218 fat that could be transported over long distances and used for pig feeding[19]. Pig-enriched
219 salmonellae exhibited greater correlation with these products than that of fresh pork and live pigs
220 (Fig. 4D-G), which also correlated to pathogen transmissions with lower, yet significant
221 coefficients of 0.46-0.72 (Fig. 5 and Extended Data Fig. 8).

222

223 **Discussion**

224 The modern agriculture system, including intensive farming and global transportation, has
225 significantly recrafted the daily life of not only livestock animals but also ourselves[6]. But how
226 much has it modified the life of bacteria, particularly, the zoonotic pathogens? Here based on
227 genomic analysis of >9000 pig-associated strains, we demonstrated that the modernization and
228 globalization of agriculture in the past century had driven the emergences, population expansions,
229 ARG acquisitions, and global transmissions of pig-enriched salmonellae.

230 We initiated the investigation in Choleraesuis, a prominent serovar that has been specifically
231 infecting both pigs and humans for >2000 years[11]. Compared to previous studies that focused on
232 either its ancient divergence or regional dissemination[11], we compiled a global dataset of
233 Choleraesuis strains to give a comprehensive overview of its recent evolution. We demonstrated

234   the high occurrences of cross-continental transmissions with a peak in the 1950s-1980s, a period

235   of accelerated globalization before the use of specific vaccines[20]. The most obvious example is

236   between the Chinese mainland and Taiwan. Despite their geographic closeness, almost all strains

237   in the Chinese mainland were in Clade 1.2 and imported from Europe, whereas strains in Chinese

238   Taiwan were in Cluster 2.2.4 and imported from the US. We attributed this to the different trading

239   partners between the two regions. The Chinese mainland mostly imported pork and pig breeds

240   such as Yorkshires and Landraces from Europe during 1950-2000[21], whilst Chinese Taiwan traded

241   more frequently with the US[22].

242   The quinolone-resistant *Salmonella* has been regarded as an "Urgent threat" and is widely found

243   in Typhi and Paratyphi A[23]. We found that 38% and 10.8% of Choleraesuis strains from humans

244   and pigs were also quinolone-resistant, mediated by either mutation in QRDR (quinolone

245   resistance-determining regions) or acquisition of ARGs (Additional Supplementary Files Table 2).

246   Furthermore, many quinolone-resistant strains, especially the human strains in Clusters 1.1.1,

247   1.2.1, and 2.2.4, also exhibit resistance to many antimicrobials extensively used in animals and

248   clinical settings and expose imminent threats to public health. For example, the strains from

249   Cluster 2.2.4 have been epidemic in Chinese Taiwan during 1995-2003[16]. Particularly, compared

250   to those from wild boar, strains from pigs and humans had $10\times$ more ARGs or QRDR mutations,

251   revealing a strong association between the extensive animal use of antimicrobials and the

252   emergence of extensive-resistant pathogens[24] and demonstrating the importance of the "One

253   Health" strategy in the control of antimicrobial overuse.

254   Nine pig-enriched *Salmonella* populations were identified from the ~362K genomes in

255   EnteroBase. Some Typhimurium clades were previously reported as pig-adapted based on small,

256   local datasets[25], but not in our survey of >70K genomes (Fig. 1C). In contrast, we showed that two

257   prominent pig-associated serovars, ceBG1272 (Choleraesuis) and ceBG3 (Derby), have been

258   associated with pigs for millennia[26,27] and experienced significant population expansions after the

259   1900s. Besides, the other seven pig-enriched populations were not associated with pigs until the

260   1900s, during which they all experienced host jumps into pigs. We noticed the development of

261   intensive pig farming during the beginning of the 1900s, which could increase the pig-pig contacts

262   while reducing those between pigs and humans. Such an environment could increase the chance of

263   transmission among the pig populations while reducing spillover between the hosts, facilitating the

264   establishment of host-enriched pathogens. Similar host transfer and population expansion have

265   been observed in *Mycobacterium tuberculosis*, which was associated with increased contact

266   between humans after the invention of fire use[28]. Thus, we attributed the increased host jumps and

267   population expansion of salmonellae in the pig population to the development of intensive pig

268   farming in the 20th century, which is in need of further research.

269   We showed that the vast majority of the contemporary non-human strains from these nine

270   populations were from pig/pork, and proved that pig is the primary source of their host transfer

271   events (Fig. 3B). The chance that strains from these nine pig-enriched populations being

272   transmitted by other non-human hosts, while could still occur, is low. Thus, we hypothesized that

273   pig-related routes, including pork and pigs, the major sources of *S. enterica* infections[29] were the

274   dominant ways for the transmission of these *salmonellae*. International salmonellosis outbreaks

275   due to global transportation of end products have been extensively reported, such as those by

276   hams[30] or chocolates[31]. However, most of these events resulted in infections in humans, which

277   represents a sink of the pathogen and rarely mediated population expansion or secondary

278   transmissions into animals[32]. In contrast, transmissions via poultry breeding stocks[33] or animal

279   feeds[30] could lead to long-term epidemics or permanent establishment of the bacteria in the target

280   regions. Furthermore, transmission of pig pathogens has been more frequently associated with the

281   end products due to the common application of swill feeding[19], as evidenced in foot-and-mouth

282   disease virus (FMDV) outbreaks[4], African swine fever virus (ASFV)[5], and trichinosis[34]. Based on

283   the UMAP analysis, we revealed strong associations between the global pork trade and the

284   transmissions of pig-enriched salmonellae. This indicated that the predominance of modern pig

285   industries in Europe and the Americas made them the centers of development and global

286   dissemination of salmonellae, highlighting the role of agricultural practice as a driver of the

287   geographic dispersal of associated bacterial pathogens.

288   A limitation of this work is that the majority of involved genomes were from public databases and

289   had sampling bias towards developed countries. The genetic diversities of pig-enriched

290   salmonellae in the majority of developing countries, especially those in South America and Africa

291   have not been sampled. Downsampling was adopted to reduce the influence of such bias, but, as

292   evidenced in Choleraesuis, itself could introduce a bias towards countries of low genetic

293   diversities. Furthermore, there is not enough data for investigation of local transmissions driven by

294   country-wide agricultural transportation or the movements of wild boars, as previously reported [35].

295   In summary, our findings demonstrate the influence of modern agriculture on the population

296   dynamics of *S. enterica*. The intensive farming has driven the host jumps of many *S. enterica*

297   populations into pigs and population expansion of the pig-associated populations, the widespread

298   availability of antibiotics after the 1940s increased the prevalence of antimicrobial resistance

299   (AMR), and the expansion of globalized trade and transportation resulted in rapid and frequent

300   global dissemination of these pig-enriched *S. enterica* (Fig. 5). Despite decades of significant

301   progress on *Salmonella* control in pigs, the evidence provided here warrants further investigation

302   and potential intervention into the global spread of *S. enterica* from centralized origins at the

303   pinnacle of pork production.

304

305   **Methods**

306   **Strains and whole-genome sequencing procedures**

307   The metadata associated with all 362,931 *S. enterica* strains accessible in EnteroBase (July 2022)

308   was downloaded and manually classified into seven categories: Pig, Bovine, Poultry, Human, other

309   Animals, and Food/Environment. The serovar associated with each ceBG (HC900 cluster) was

310   downloaded from https://enterobase.readthedocs.io/en/latest/HierCC_lookup.html. A subset of 61

311   ceBGs with ≥20 pig strains (277,588 strains in total) was used to display as a hierarchical bubble

312   plot (Fig. 1D), showing phylogenetic grouping at subspecies (HC2850), super-lineage (HC2000), and

313   ceBG (HC900) levels, along with pie charts representing source category. A total of 9,259 genomes

314   from 9 ceBGs each containing >40% of pig-associated strains were selected for downstream

315   analysis. Additionally, a set of 16,829 genomes from 8 pig-containing ceBGs that each contain

316   lower levels (2-40%) of pig-associated strains were also selected in comparison with the 9 pig-

317   enriched ceBGs. Additionally, 15 Choleraesuis strains were collected by China CDC from northern

318   and eastern regions across China between 2002 and 2022. The DNA of each strain was extracted

319   using the HiPure Bacterial DNA kit (D3146). Paired-end libraries with insert sizes of ~300 bp were

320   prepared following Illumina's standard genomic DNA library preparation procedure (VAHTS

321   Universal DNA Library Prep kit for Illumina V3) and sequenced on an Illumina NovaSeq 6000

322   using the S4 reagent kits (v1.5) according to the manufacturer's instructions.

323   To demonstrate the association between pig and its enriched *S. enterica* strains, we sequenced the

324   genome of 78 strains of the prominent pig-enriched serovar, Choleraesuis. These include 63

325   strains from Germany or Austria as part of the University of Warwick/University College Cork

326   10K genomes project[36] and 15 from China. They were integrated with public genomes of 679

327   strains isolated between 1935 and 2022 and one genotype reconstructed from a ~1600-year-old

328   human remains[37], resulting in a global collection of 757 genomes (Additional Supplementary Files

329   Table 2) encompassing 41 countries.

330

331   **Bioinformatic analysis**

332   The sequencing reads of each strain were quality trimmed using EtoKi prepare[12], and the high-

333   quality sequences were further assembled into contigs using SPAdes V3.13[38] which was

334   implemented in the 'EtoKi assemble' pipeline. The genes in each assembled genome were predicted

335   and annotated using PROKKA 1.14.6[39] and had detailed functional predictions using eggnog-

336   mapper v2[40]. The antibiotic resistance genes were predicted using AMRfinder v3.11.14[41], and the

337   disrupted genes in the pig-enriched populations were predicted using PEPPAN[42]. The multi-

338   sequence alignment for each pig-enriched ceBG was generated using the EtoKi align module and

339   used to build a maximum-likelihood (ML) phylogeny using IQTree v1.6.12[43] implemented in EtoKi

340   phylo after the removal of recombinant regions using RecHMM[11].

341

342   **Temporal signal and randomization test**

343   The presence of a temporal signal in *Salmonella* serovar Choleraesuis (ceBG1272) was tested using

344   three approaches. The regression of root-to-tip distances and dates of isolation was estimated using

345   TempEst v1.5.3[44] with a correlation of determination (R^2) of 0.67 and P-value of 4.15×10^{-6}. We

346   then randomly permutated the isolation dates of the strains ten times and estimated their R^2 values.

347   The same datasets were also used for BactDating v1.1[45] inferences as described above, and their

348   substitution rates were compared with the rate from the actual data (Extended Data Fig. 2),

349 demonstrating the presence of a significant temporal signal. We also performed the same tests

350 without the ancient genotype, demonstrating the presence of temporal signals. Furthermore, the

351 tests were also performed for the other eight pig-enriched populations, proving their availability for

352 temporal analyses (Extended Data Fig. 3).

353

Population dynamics of serovar Choleraesuis

355 The ML tree of serovar Choleraesuis was calibrated by dated tips using BEAST[44] with a GTR

356 substitution model and fixed topology. Eight BEAST runs were prepared by combinations of two

357 clock models of "strict clock rate" and "optimised relaxed clock", and four population models of

358 "constant coalescence", "Bayesian skyline", "Birth-death skyline" and "extended Bayesian

359 skyline". All models were run in "Nested Sampling" mode with 8 parallel chains each with

360 "chainLength=20000", "particleCount=1", and "subChainLength=5000". The results were

361 summarized using NSLogAnalyser and the model with "optimised relaxed clock" and "Bayesian

362 skyline" had the greatest marginal likelihood with a maximum ESS of 632.3. The posterior trees

363 from the best model were then summarized using treeannotator into a maximum clade credibility

364 tree (MCC tree) and visualized in iTol[46].

365 Two downsampling tests were performed by selecting at most ten or five random genomes from

366 each country. We employed TreeTime to reconstruct the ancestral states of the internal nodes based

367 on a subtree containing only those selected tips. Each test was run parallel 100 times, and the results

368 were summarized in Extended Data Fig. 4.

369

Population dynamics of pig-enriched and pig-containing ceBGs

371 Furthermore, the population dynamics of each pig-enriched population were estimated using

372 BactDating[47], which performed Bayesian inference of ancestral dates based on the ML tree. Parallel

373 chains of 5e6 samples each were run for each of the substitution models of "strictgamma",

374 "mixedgamma", and "carc". The first 50% of the chain (3e6 samples) for each model was discarded

375 as burn-ins and the convergence of the run was determined by ensuring effective sampling sizes

376 (ESSs) of >100 for all parameters. The results from all samples were compared based on their

377 Bayes Factors using the modelcompare function in BactDating, and only the best model for each

378   pig-enriched population was reported. Notably, the dating results for ceBG1272 (Choleraesuis) by

379   BactDating were very similar to that by BEAST, suggesting high reproducibility of the analyses.

380   We then estimated the host transfers and geographic transmissions of each population along the

381   dated trees using the ML algorithm implemented in TreeTime[48]. Similarly, the dates and host

382   transfers of the ancestral nodes in the ML trees of the pig-containing ceBGs were also estimated as

383   described above.

384

Transmission events and the correlation between trade data

386   We define a transmission or host transfer event when the ancestral node and the descending node

387   of a branch are assigned different country/host information. Considering all possible states as S.

388   A host transfer or international transmission was counted along the ML tree if the reconstructed

389   states in the ancestral and the descending nodes of a branch were different, and the numbers of

390   host transfers and international transmissions were summarized as $N_{i->j}$, where $i \in S$ and $j \in S$ are

391   the states of the ancestral and descending nodes, respectively. Then there is:

$$T_{i \to j} = N_{i \to j} \bigg/ \sum_{k \in S} N_{k \to j}$$

393   Where $T_{i->j}$ is the normalized frequency of Sj originating from Si, and $k \in S$ is an iterator for

394   calculating the total number of transmissions into Sj. Furthermore, the normalized frequency

395   between two continents m and n are:

$$\hat{T}_{m \to n} = \sum_{a \in m} \sum_{b \in n} T_{a \to b}$$

397   Where a and b are countries in the continents m and n, respectively.

398

399   The international trade data of all categories were obtained from Harvard Dataverse

400   (https://dataverse.harvard.edu). Trade values were expressed in constant U.S. dollars after

401   adjustment for inflation and summed up. The normalized fluxes of trade between continents were

402   also calculated using a procedure similar to the above. As a result, frequencies of pathogen

403   transmissions and trading among five continents of Asia, Africa, Europe, Oceania, and the

404 Americas were obtained and pairwise Pearson's correlation coefficients (R) between the cross-
405 continental flows of the goods and between the flows of the goods and the pathogens were
406 calculated. All goods and salmonellae were then projected into a 2-D space based on their *1-R*
407 values using UMAP (Uniform Manifold Approximation and Projection), which is a non-linear
408 dimension reduction technique that has been extensively used in biological analysis, such as in
409 single-cell studies[49].

410

Data availability

412 The raw sequencing reads for the 15 Chinese strains have been deposited in the Genome Sequence
413 Archive in the National Genomics Data Center, China National Center for Bioinformation / Beijing
414 Institute of Genomics, Chinese Academy of Sciences (GSA: CRA012579) and are publicly
415 accessible at https://ngdc.cncb.ac.cn/gsa. The assembled genome sequences have been deposited in
416 the Genome Warehouse (GWH) in the National Genomics Data Center with BioProject accession
417 PRJCA019682. The raw reads for 67 European Choleraesuis strains were deposited in Short Reads
418 Archive (SRA) at EBI under BioProject accession: PRJEB20997, as part of the University of
419 Warwick/University College Cork (UOWUCC) 10K genomes project. A detailed list of the sample
420 accession codes for all Choleraesuis strains is available in Additional Supplementary Files Table 2.
421 Assembled genomes for all pig-enriched populations were available as a workspace in EnteroBase
422 at https://enterobase.warwick.ac.uk/a/100355. The resulting figures and underlying data of 61
423 ceBGs with ≥20 pig strains are all available at https://observablehq.com/d/232a986be1a99113.

431 **Competing interests**

432 The authors declare no conflicts of interest.

433 **Authors' contributions**

434 HL, YR, XL, ZZ designed the study. UM, XL applied the experiments, DFeng, QJ, JR, and LB
435 prepared the figures and tables. LZ, SL, YR, ZL analysed the statistical data. HL, YW, DFeng, ZZ
436 wrote the initial version of the manuscript. HL, DFeng, SO, DFalush, XD, BK, ZZ revised the
437 manuscript.

438

439

440   **Figure Legends**

441   **Fig. 1 | Summary of the pig-enriched ceBGs in the *Salmonella* database in EnteroBase.** (A)
442   Histogram of the numbers of *Salmonella* strains for each non-human source in EnteroBase. (B)
443   Histogram of the numbers of pig-associated *Salmonella* strains per year during 1885-2022. (C) Bubble
444   plot of the 61 pig-associated ceBGs each with ⩾20 strains. Each ceBG is sized proportional to the
445   number of strains in it, and placed according to the numbers and percentages of its pig-associated
446   strains. (D) Hierarchical bubble plot of the 61 pig-associated ceBGs as in panel C. The sizes of the
447   circles are proportional to the number of strains, the three levels in the plot represent (from outer to
448   inner) clusters at the levels of subspecies (HC2850), super-lineage (HC2000), and ceBG (HC900),
449   as described previously[9]. Pie charts represent the proportions of strains from different sources. The 9
450   pig-enriched ceBGs each with >40% pig strains were labelled in parts C and D.

451   **Fig. 2 | Population dynamics, ARGs, and global transmissions of ceBG1272 (Choleraesuis).** (A) The
452   maximum-likelihood phylogeny (left), metadata (middle), and predicted ARGs for all strains in
453   ceBG1272. The predicted lineages, clades, and clusters are labeled near the associated branches. The
454   associations between the numbers of predicted ARGs and the sampling years, source categories, and
455   countries are also visualized. (B) The associations between the numbers of predicted ARGs and the
456   sampling years in Para C lineages, and (C) the association between the number of ARGs per strain and
457   host sources. (D) Visualization of the correlations between the numbers of predicted ARGs and countries
458   in Choleraesuis. (E, F, and G) Bayesian inferences of the population dynamics of ceBG1272 over the past
459   ~2500 years. (E) Global transmission of ceBG1272. Pie charts show the proportional composition of
460   clades in each major country, and the arrows show the transmissions reconstructed based on the tree in
461   part F, with the transmission dates shown nearby. The world map was modified from the map hosted in
462   the d3.js (https://d3js.org/). The piecharts and arrows were color-coded by the associated clades.
463   Inset: the Simpson diversity of clades in each geographic region in the world. (F) The maximum clade-
464   credibility (MCC) tree of ceBG1272 by BEAST2. The branches were color-coded by the most probable
465   ancestral geographic origins (as in the Key). Piecharts of all possible geographic origins were shown over
466   certain nodes where the most probable origins had <90% posterior supports. The dates of origin for some
467   branches were shown together with the 95% confidence intervals in brackets. (G) The fluctuation of
468   effective population sizes with time by the 'skygrowth' package in R. Arrows points to the time of three
469   major developments in the modern pig industry.

470   **Fig. 3 | Host transfers for the pig-enriched ceBGs.** (A) Curves show the dynamic changes of the
471   proportional host sources with time for each ceBG. The ancestral host associations were predicted by

472   TreeTime. The predicted median effective population sizes were also shown as black curves for

473   ceBG1272 (Choleraesuis) and ceBG3 (Derby). The period for host transfers into pigs (red) or population

474   expansions (yellow) is shown above each plot. Detailed prediction of population dynamics for all nine

475   ceBGs can also be found in Extended Data Fig. 5. (B) Proportional source of host transfers summarized

476   for all nine ceBGs in the past 50 years. Detailed host transfer data for each ceBG can be found in

477   Extended Data Fig. 6. (C) Proportional source of host transfers in the past 50 years summarized for eight

478   pig-containing ceBGs that have 5-35% pig strains, including ceBGs of 5, 8, 22, 125, 276, 709, and 1898

479   (see Fig. 1C). (B, C) The arrows show the direction of transfers and are color-coded by average

480   frequencies. The most contributing sources (B: pigs, and C: human) are highlighted with a '*'. (D)

481   Median numbers of disrupted CDSs (coding sequences) per genome with 95% confidence intervals in all

482   pig-enriched and pig-containing ceBGs.

483   **Fig. 4 | Association between the transmission of the pig-enriched ceBGs and the global trade of pig-**

484   **related products.** (A, B) Visualization of the international trade of all pork-related products (top) and the

485   transmissions of pig-enriched ceBGs (bottom). The pie charts show the relative proportions of source

486   continents for the products or pathogens to the target continents. The world map was modified from the

487   map hosted in the d3.js (https://d3js.org/). (C) UMAP plot of Pearson's correlations among the trade and

488   the transmission of pig-enriched salmonellae. Each colored dot in the plot shows an animal-related

489   product as in the Harvard database and the grey dots are other, non-animal products. The triangle shows

490   inter-continental transmission data summarized from all pig-enriched salmonellae. The insert highlights

491   the dashed box in the plot, with arrows specifying the correlation coefficient (R) between the

492   transmissions of pathogens and the trade of pig-related products. (D-G) Linear regressions of different

493   categories of pig-related products (X-axis) and the inter-continental transmissions of pig-enriched ceBGs

494   (Y-axis). The correlation coefficient values for the linear regressions are also shown.

495   **Fig. 5 | The influences of modern agriculture on population dynamics of *Salmonella enterica***

496   **serovars.** This image provides a visual summary of the paper. Agricultural production has become

497   increasingly modernized over the past half-century. On the one hand, the pattern of large-scale intensive

498   pig farming has led to the emergence and population expansion of pig-enriched *Salmonella*; on the other

499   hand, globalized trade exchanges concerning pigs have similarly increased the probability of global

500   transmission of pig-enriched *Salmonella enterica* serovars. In addition, with the use of antibiotics in the

501   process, more and more pig-enriched *Salmonella* has obtained new antibiotic resistance genes, including

502   many of the previously reported human-specific antibiotic resistance genes. The impact of the

503   development model of modern agriculture on pig-enriched *Salmonella* is comprehensive and far-reaching.

504

References

1. López-Gálvez, F., Gómez, P. A., Artés, F., Artés-Hernández, F. & Aguayo, E. Interactions between Microbial Food Safety and Environmental Sustainability in the Fresh Produce Supply Chain. *Foods* **10**, 1655 (2021).

2. Stanaway, J. D. *et al.* The global burden of typhoid and paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Infect Dis* **19**, 369–381 (2019).

3. Bonardi, S. *Salmonella* in the pork production chain and its impact on human health in the European Union. *Epidemiol. Infect.* **145**, 1513–1526 (2017).

4. Jamal, S. M. & Belsham, G. J. Foot-and-mouth disease: past, present, and future. *Vet Res* **44**, 116 (2013).

5. Wu, X., Fan, X., Xu, T. & Li, J. Emergency preparedness and response to African swine fever in the People's Republic of China: -EN- -FR- Préparation et réponse à la crise de peste porcine africaine en République populaire de Chine -ES- Preparación para emergencias y respuesta a la peste porcina africana en la República Popular de China. *Rev. Sci. Tech. OIE* **39**, 591–598 (2020).

6. Woods, A. Rethinking the History of Modern Agriculture: British Pig Production, c.1910-65. *20 Century Br Hist.* **23**, 165–191 (2012).

7. Bosse, M. *et al.* Genomic analysis reveals selection for Asian genes in European pigs following human-mediated introgression. *Nat Commun* **5**, 4392 (2014).

8. Koopman, R. B., Laney, K. & Giamalva, J. Industry and Trade Summary. Publication ITS-011. Washington, DC: U.S. International Trade Commission (2014).

9. Achtman, M., Zhou, Z., Charlesworth, J. & Baxter, L. EnteroBase: hierarchical clustering of 100 000s of bacterial genomes into species/subspecies and populations. *Phil. Trans. R. Soc. B* **377**, 20210240 (2022).

10. Gal-Mor, O. Persistent Infection and Long-Term Carriage of Typhoidal and Nontyphoidal Salmonellae. *Clin Microbiol Rev* **32**, e00088-18 (2018).

11. Zhou, Z. *et al.* Pan-genome Analysis of Ancient and Modern *Salmonella enterica* Demonstrates Genomic Stability of the Invasive Para C Lineage for Millennia. *Curr Biol* **28**, 2420-2428.e10 (2018).

12. Zhou, Z. *et al.* The EnteroBase user's guide, with case studies on *Salmonella* transmissions, *Yersinia pestis* phylogeny, and *Escherichia* core genomic diversity. *Genome Res.* **30**, 138–152 (2020).

13. Wang, R. *et al.* The global distribution and spread of the mobilized colistin resistance gene mcr-1. *Nat Commun* **9**, 1179 (2018).

14. Shambaugh, G. E. History of Sulfonamides. *Arch Otolaryngol Head Neck Surg.* **83**, 1–2 (1966).

15. Chase-Dunn, C., Kawano, Y. & Brewer, B. D. Trade Globalization since 1795: Waves of Integration in the World-System. *Am. Sociol. Rev.* **65**, 77 (2000).

539   16. Chiu, C.-H., Su, L.-H. & Chu, C. *Salmonella enterica* Serotype Choleraesuis: Epidemiology,
540   Pathogenesis, Clinical Disease, and Treatment. *Clin Microbiol Rev* **17**, 311–322 (2004).

541   17. Massei, G. *et al.* Wild boar populations up, numbers of hunters down? A review of trends and
542   implications for Europe: wild boar and hunter trends in Europe. *Pest. Manag. Sci.* **71**, 492–500
543   (2015).

544   18. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for
545   Dimension Reduction. Preprint at http://arxiv.org/abs/1802.03426 (2020).

546   19. Dame-Korevaar, A., Boumans, I. J. M. M., Antonis, A. F. G., Van Klink, E. & De Olde, E. M.
547   Microbial health hazards of recycling food waste as animal feed. *Future Foods* **4**, 100062 (2021).

548   20. Aparicio, G., González-Esteban, Á. L., Pinilla, V. & Serrano, R. The World Periphery in Global
549   Agricultural and Food Trade, 1900–2000. *Agricultural Development in the World Periphery* 63–88
550   (2018).

551   21. Gale, F., Marti, D. & Hu, D. China's Volatile Pork Industry. *Economic research service report,*
552   *United States Department of Agriculture*, pp. 1-27(2012).

553   22. Simon, S. Real People, Real Dogs, and Pigs for the Ancestors: The Moral Universe of
554   "Domestication" in Indigenous Taiwan: Real People, Real Dogs, and Pigs for the Ancestors. *Am*
555   *Anthropol.* **117**, 693–709 (2015).

556   23. Luhmann, N., Holley, G. & Achtman, M. BlastFrost: fast querying of 100,000s of bacterial genomes
557   in Bifrost graphs. *Genome Biol* **22**, 30 (2021).

558   24. Chiu, C.-H. *et al.* The Emergence in Taiwan of Fluoroquinolone Resistance in *Salmonella enterica*
559   Serotype Choleraesuis. *N Engl J Med* **346**, 413–419 (2002).

560   25. Bawn, M. *et al.* Evolution of *Salmonella enterica* serotype Typhimurium driven by anthropogenic
561   selection and niche adaptation. *PLoS Genet* **16**, e1008850 (2020).

562   26. Key, F. M. *et al.* Emergence of human-adapted *Salmonella enterica* is linked to the Neolithization
563   process. *Nat Ecol Evol* **4**, 324–333 (2020).

564   27. Sévellec, Y. *et al.* Polyphyletic Nature of *Salmonella enterica* Serotype Derby and Lineage-Specific
565   Host-Association Revealed by Genome-Wide Analysis. *Front. Microbiol.* **9**, 891 (2018).

566   28. Comas, I. *et al.* Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis
567   with modern humans. *Nat Genet* **45**, 1176–1182 (2013).

568   29. Roasto, M. *et al. Salmonella enterica* prevalence, serotype diversity, antimicrobial resistance and
569   control in the European pork production chain. *Trends Food Sci Technol* **131**, 210–219 (2023).

570   30. Dawoud, T. M., Shi, Z., Kwon, Y. M. & Ricke, S. C. Overview of Salmonellosis and Food-borne
571   *Salmonella*. *Producing Safe Eggs* 113–138 (2017).

572   31. Werber, D. *et al.* International outbreak of *Salmonella* Oranienburg due to German chocolate. *BMC*
573   *Infect Dis* **5**, 7 (2005).

574   32. Sokurenko, E. V., Gomulkiewicz, R. & Dykhuizen, D. E. Source–sink dynamics of virulence
575   evolution. *Nat Rev Microbiol* **4**, 548–555 (2006).

576   33. Li, S., He, Y., Mann, D. A. & Deng, X. Global spread of *Salmonella* Enteritidis via centralized
577   sourcing and international trade of poultry breeding stocks. *Nat Commun* **12**, 5109 (2021).

578   34. Foreyt, W. J. *Trichinosis: Reston, Va., U.S. Geological Survey Circular 1388, 60 p., 2 Appendixes*.
579   (2013).

580   35. Leekitcharoenphon, P. *et al.* Cross-Border Transmission of *Salmonella* Choleraesuis var. Kunzendorf
581   in European Pigs and Wild Boar: Infection, Genetics, and Evolution. *Front. Microbiol.* **10**, 179
582   (2019).

583   36. Achtman, M. *et al.* Genomic diversity of *Salmonella enterica* -The UoWUCC 10K genomes project.
584   *Wellcome Open Res* **5**, 223 (2021).

585   37. Key, F. M. *et al.* Emergence of human-adapted *Salmonella enterica* is linked to the Neolithization
586   process. *Nat Ecol Evol* **4**, 324–333 (2020).

587   38. Bankevich, A. *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-
588   Cell Sequencing. *J Comput Biol* **19**, 455–477 (2012).

589   39. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).

590   40. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper
591   v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic
592   Scale. *Mol Biol Evol* **38**, 5825–5829 (2021).

593   41. Bortolaia, V. *et al.* ResFinder 4.0 for predictions of phenotypes from genotypes. *J Antimicrob*
594   *Chemother* **75**, 3491–3500 (2020).

595   42. Zhou, Z., Charlesworth, J. & Achtman, M. Accurate reconstruction of bacterial pan- and core
596   genomes with PEPPAN. *Genome Res.* **30**, 1667–1679 (2020).

597   43. Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: A Fast and Effective
598   Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol* **32**, 268–274
599   (2015).

600   44. Didelot, X., Croucher, N. J., Bentley, S. D., Harris, S. R. & Wilson, D. J. Bayesian inference of
601   ancestral dates on bacterial phylogenetic trees. *Nucleic Acids Res.* **46**, e134–e134 (2018).

602   45. Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure of
603   heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* **2**, vew007 (2016).

604   46. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments.
605   *Nucleic Acids Res.* **47**, W256–W259 (2019).

606   47. Bouckaert, R. *et al.* BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS*
607   *Comput Biol* **10**, e1003537 (2014).

608   48. Sagulenko, P., Puller, V. & Neher, R. A. TreeTime: Maximum-likelihood phylodynamic analysis.
609   *Virus Evol* **4**, (2018).

610   49. Becht, E. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* **37**,
611   38–44 (2019).

612

613

614

Extended Data Figure 1. Circular presentation of the maximum-likelihood phylogeny in Figure 1A.
Outer rings: The clade, source, and geographic origin of each strain. Colored arcs underneath the tree
show the cluster assignments as in the Key.

Extended Data Figure 2. Evaluation of the presence of temporal signal in *S. enterica* serovar
Choleraesuis. The analysis was performed on 587 genomes. (A) Linear regression between root-to-tip
distances of strains and the sampling years with a coefficient of determination (R^2) of 0.67. (B)
Substantially lower R^2 values were obtained for ten date-randomisation datasets. (C) The average (dots)
and standard deviation (error bars) of the substitution rates for actual data (black) and ten date-
randomisation datasets (red), estimated by BactDating.

Extended Data Figure 3. Assessing the existence of temporal signal randomization test in pig-
enriched ceBGs. (A) Coefficients of determination (R^2) for the ten date-randomisation tests were
obtained by linear regression between root-to-tip distances of strains and the sampling years in pig-
enriched ceBGs. (B) The average (dots) and standard deviations (error bars) of the substitution rates for
actual data (black) and ten date-randomisation datasets (red) by BactDating. Pig-enriched ceBGs
including ceBG3 (n=3136), ceBG10 (n=622), ceBG17 (n=155), ceBG35 (n=516), ceBG37 (n=1012),
ceBG40 (n=176), ceBG459 (n=1441), ceBG621 (n=787), ceBG1272 (n=586). The ancient sample was
not included in ceBG1272 in both (A) and (B).

Extended Data Figure 4. The geographic states of the ancestral nodes in the Choleraesuis
population after downsamplings. (A) Downsampling tests of up to ten strains per country/region. (B)
Downsampling tests of up to five strains per country/region. Both: the geographic states were predicted
using TreeTime. Each test was run in 100 parallels. The pie charts illustrate the proportions of the best-
supported geographic states in the 100 parallels for the corresponding ancestral nodes.

Extended Data Figure 5. Reconstructed ancestral host associations for all nodes in nine pig-
enriched ceBGs. The corresponding ceBGs for each panel are: (A) ceBG3 (Derby), (B) ceBG1272
(Choleraesuis), (C) ceBG17 (Chailey), (D) ceBG35 (Worthington), (E) ceBG37 (London), (F) ceBG40
(Cerro), (G) ceBG459 (Johannesburg), (H) ceBG621 (Ohio), (I) ceBG10 (Adelaide). The mean values of
effective population sizes with time were also shown for (A) and (B), with 95% confidence intervals in
grey shapes. The orange and grey boxes show the periods of population expansion (A, B) or the periods
of host transfers into pigs (C-I).

Extended Data Figure 6. The host transmission frequency of nine pig-enriched ceBGs. The
corresponding ceBGs for each panel are: (A) ceBG10 (Adelaide), (B) ceBG1272 (Choleraesuis), (C)

646   ceBG17 (Chailey), (D) ceBG3 (Derby), (E) ceBG35 (Worthington), (F) ceBG37 (London), (G) ceBG 40

647   (Cerro), (H) ceBG 459 (Johannesburg), (I) ceBG 621 (Ohio). Different arrows represent the direction of

648   host transmission, with darker colors indicating higher transfer frequency. "*" marks the most

649   contributing host sources for the transmissions.

650   **Extended Data Figure 7. A 5 X 5 table showing the summarised international transmission events**

651   **of all nine pig-enriched ceBGs in *Salmonella* in the past 50 years.** The maximum-likelihood phylogeny

652   was reconstructed based on SNPs in the core genome of each ceBG and the date of origin was estimated

653   using BactDating. TreeTime was applied to reconstruct the country sources for all nodes in the tree. A

654   transmission was recorded when the ancestral node and descending node of a branch were different. All

655   transmissions were then summarized and grouped based on their associated continents.

656   **Extended Data Figure 8. Pearson's correlations analysis of sub-classification for pig-related**

657   **products.** The Sub classifications involved include (A) Pork offal, frozen (021011); (B) Pork, frozen

658   (020649); (C) Pork, frozen cut (020329); (D) Pork, preserved (020322); (E) Pork offal, fresh (020630);

659   (F) Pig fat (020910); (G) Pork, fresh cuted (020912); (H) Pork, fresh (020319); (I) Pig carcasses

660   (020310); (J) Pig, live breeding (010391); (K) Pig, live, less 50kg (010392); (L) Pig, live, over 50kg

661   (020311). The codes in parentheses represent the Harmonized System Codes of the products. Analyzing

662   the correlation between the intercontinental dissemination of each pig-related commodity and the

663   intercontinental transmission of pig-enriched ceBGs. R: Pearson's correlation coefficient.

664

665   **Supplementary Information and Additional Supplementary Files**

666

667   **Supplementary Table S1. Pig-enriched HC5s that were isolated from multiple countries.**

668   The item in the table including the HC5 cluster from Enterobase, corresponded with pig-enriched ceBGs,

669   continent, transnational, and host sources information.

670

671   **Supplementary Table S2. Simpson's diversity of clades at the national level.**

672   When calculating the Simpson's diversity of clades, only countries with more than 10 strains

673   were selected.

674

675   **Supplementary Table S3. Distribution of the colistin-resistant genes in serovars Choleraesuis.**

676   Information on colistin-resistant genes extracted from the statistical information on the number of ARGs

677   of strains of serovar Choleraesuis is summarised in this table. The collection years, host sources, and

678   geographic information for strains with colistin resistance are labeled in the table.

679

Supplementary Table S4. The frequencies of transmissions in each pair of regions across all 100 random downsamplings.
The table on the left shows the results of 5 samples drawn from each region, while the table on the right shows the results of 10 samples drawn from each region. The transmission frequency corresponds to the total number of transmission events that occur in 100 repeats.

685

Additional Supplementary Files Table 1 Metadata of eight pig-enriched ceBGs except 1272 (Choleraesuis).
The metadata included eight pig-enriched ceBGs (which had ⩾40% of the strains from pigs) except ceBG1272 (Choleraesuis), including ceBG3 (Derby), ceBG10 (Adelaide), ceBG459 (Johannesburg), ceBG37 (London), ceBG621 (Ohio), ceBG40 (Cerro), ceBG35 (Worthington), and ceBG17 (Chailey). The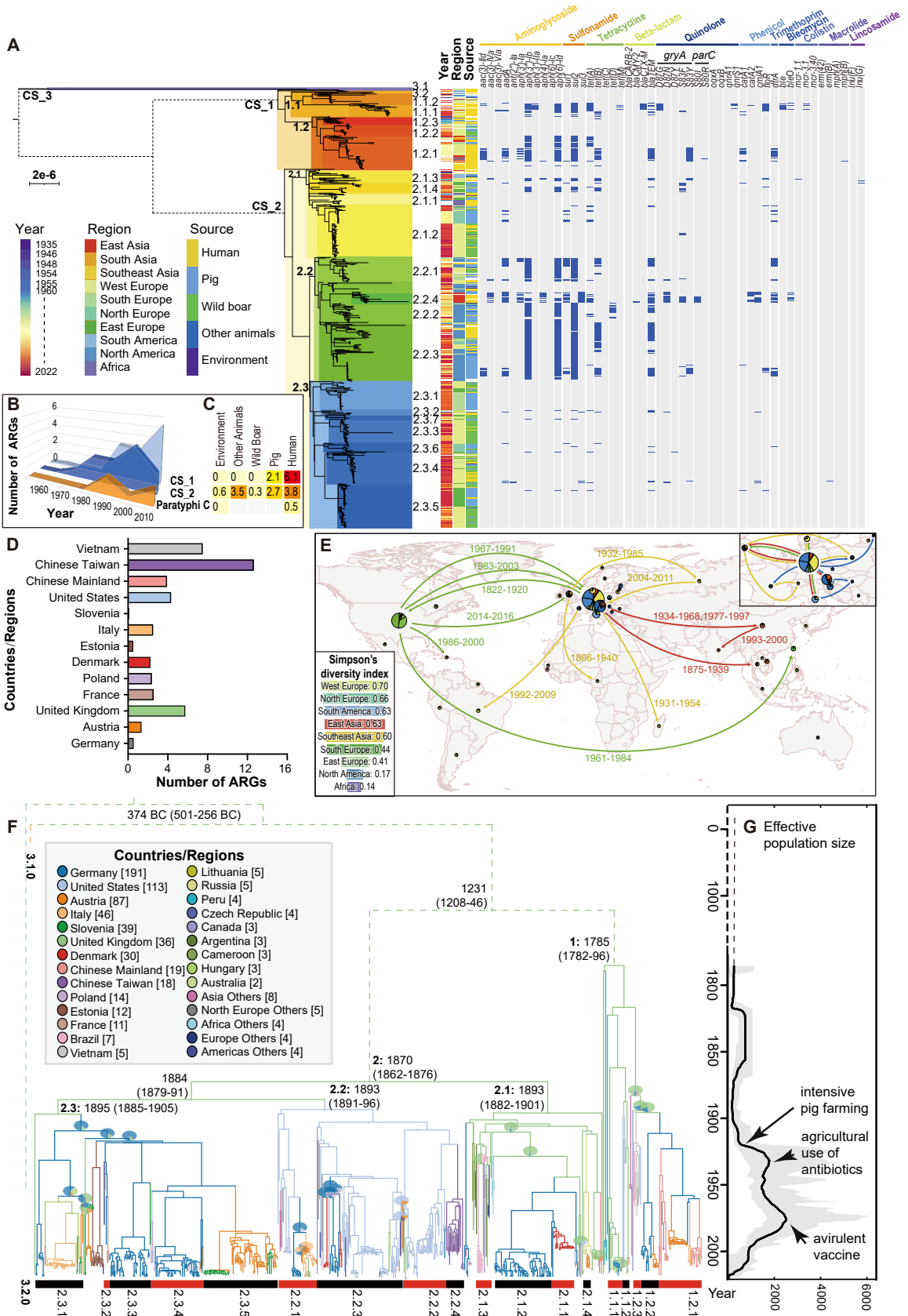 table provides information about the ceBG names, geographic information, collection year, source details, and the number of ARGs per strain of these pig-enriched ceBGs.
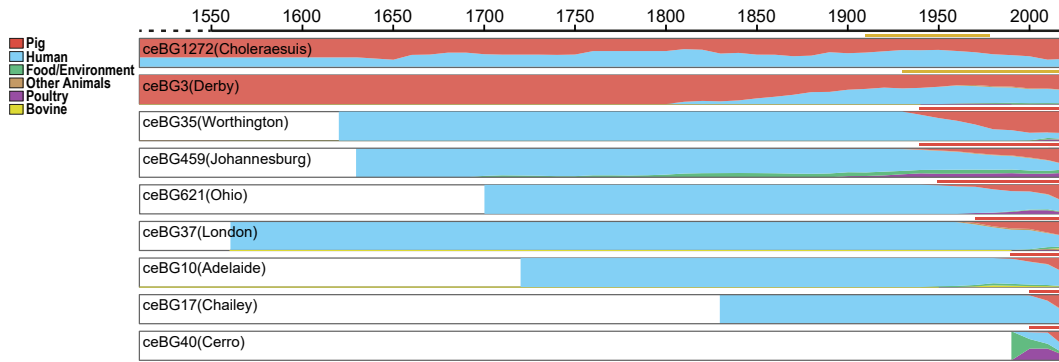
693

Additional Supplementary Files Table 2 Metadata of *Salmonella* strains in the serovars Choleraesuis and Paratyphi C. There is a total of 911 *Salmonella* strains from the Para C lineage, including 757 *S*. Choleraesuis serovars and 154 Paratyphi C serovars. The specific serotyping, geographic information, collection year, and host source details of these strains are listed in this table, as well as the statistical information about the number of ARGs per strain.
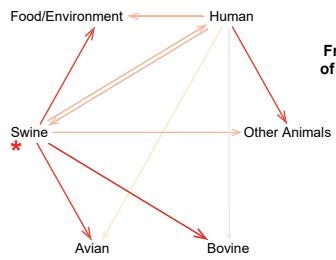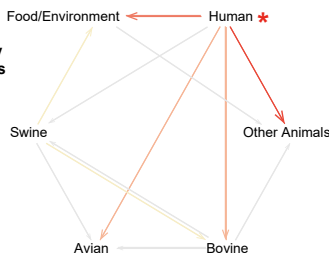
A

Number of non-human strains (y-axis): 50k, 40k, 30k, 20k, 10k, 0

x-axis categories: Poultry, Pig, Bovine, Other Animals, Water, Plant, Soil

B

Number of pig-associated strains (y-axis): 10000, 1000, 100, 10, 1

Sampling years (x-axis): 1880, 1900, 1920, 1940, 1960, 1980, 2000, 2020

C

Percentage of pig-associated strains (y-axis): 60%, 50%, 40%, 30%, 20%, 10%, 0%

Number of pig-associated strains (x-axis): 1000, 2000, 3000, 4000

Labels: 1272 (Choleraesuis), 17 (Chailey), 10 (Adelaide), 459 (Johannesburg), 40 (Cerro), 37 (London), 35 (Worthington), 3 (Derby), 621 (Ohio), 1898, 5 (Anatum/Eko), 276, 22 (Manhattan), 8 (Saintpaul), 28 (Agona), 125 (Berta), 36 (Infantis), 709 (Litchfield), 2 (Typhimurium)

D

621 (Ohio): 43%   1272 (Choleraesuis): 61%
17 (Chailey): 52%
35 (Worthington): 45%

Legend:
Pig
Human
Food/Environment
Other Animals
Poultry
Bovine
ND

10 (Adelaide): 53%
40 (Cerro): 46%
459 (Johannesburg): 52%
37 (London): 50%
3 (Derby): 46%

A

Year
1935
1946
1948
1954
1955
1960
2022

Region
East Asia
South Asia
Southeast Asia
West Europe
South Europe
North Europe
East Europe
South America
North America
Africa

Source
Human
Pig
Wild boar
Other animals
Environment

B

C

	Environment	Other Animals	Wild Boar	Pig	Human
CS_1	0	0	0	2.1	6.1
CS_2	0.6	3.5	0.3	2.7	3.8
Paratyphi C	0	0	0	0	0.5

D

Vietnam
Chinese Taiwan
Chinese Mainland
United States
Slovenia
Italy
Estonia
Denmark
Poland
France
United Kingdom
Austria
Germany

Number of ARGs

E

Simpson's diversity index
West Europe: 0.70
North Europe: 0.66
South America: 0.63
East Asia: 0.63
Southeast Asia: 0.60
South Europe: 0.44
East Europe: 0.41
North America: 0.17
Africa: 0.14

F

374 BC (501-256 BC)

3.1.0

Countries/Regions

Germany [191]	Lithuania [5]
United States [113]	Russia [5]
Austria [87]	Peru [4]
Italy [46]	Czech Republic [4]
Slovenia [39]	Canada [3]
United Kingdom [36]	Argentina [3]
Denmark [30]	Cameroon [3]
Chinese Mainland [19]	Hungary [3]
Chinese Taiwan [18]	Australia [2]
Poland [14]	Asia Others [8]
Estonia [12]	North Europe Others [5]
France [11]	Africa Others [4]
Brazil [7]	Europe Others [4]
Vietnam [5]	Americas Others [4]

1231
(1208-46)

1: 1785
(1782-96)

1884
(1879-91)

2: 1870
(1862-1876)

2.2: 1893
(1891-96)

2.1: 1893
(1882-1901)

2.3: 1895 (1885-1905)

3.2.0

2.3.1 2.3.2 2.3.3 2.3.4 2.3.5 2.2.1 2.2.3 2.2.2 2.2.4 2.1.3 2.1.2 2.1.1 2.1.4 1.1.1 1.1.2 1.2.2 1.1.3 1.2.1

G Effective population size

intensive pig farming
agricultural use of antibiotics
avirulent vaccine

Year

A Predicted host association with time

Legend:
- Pig
- Human
- Food/Environment
- Other Animals
- Poultry
- Bovine

ceBG1272(Choleraesuis)
ceBG3(Derby)
ceBG35(Worthington)
ceBG459(Johannesburg)
ceBG621(Ohio)
ceBG37(London)
ceBG10(Adelaide)
ceBG17(Chailey)
ceBG40(Cerro)

B Pig-enriched ceBGs

Food/Environment — Human
Swine *
Other Animals
Avian — Bovine

Frequency of transfers
0.1 0.2 0.3 0.4 0.5 0.6

C Pig-containing ceBGs

Food/Environment — Human *
Swine
Other Animals
Avian — Bovine

D

Number of CDS disruptions

— Median
CI 95%

ceBGs: 1272 459 28 5 125 621 10 3 35 276 40 709 1898 8 17 22 37

A Pork-related products

$3B-5B
$200M-3B
$0-200M

Size
5
4
3
2
1
0

Source
Asia
Africa
Europe
Americas
Oceania

B Pig-enriched ceBGs

40-70
10-40
0-10

C

UMAP2

.87
.89
.96
.82
.87
.93

Pathogen
Pig
Bovine
Poultry
Non-animals

UMAP1

D Pork frozen

R=0.95

Transmission sources

Product sources

E Pork processed

R=0.89

Transmission sources

Product sources

F Pork fresh

R=0.70

Transmission sources

Product sources

G Pig live

R=0.49

Transmission sources

Product sources

Modern Agriculture

Intensive Pig Farming

Host jumps

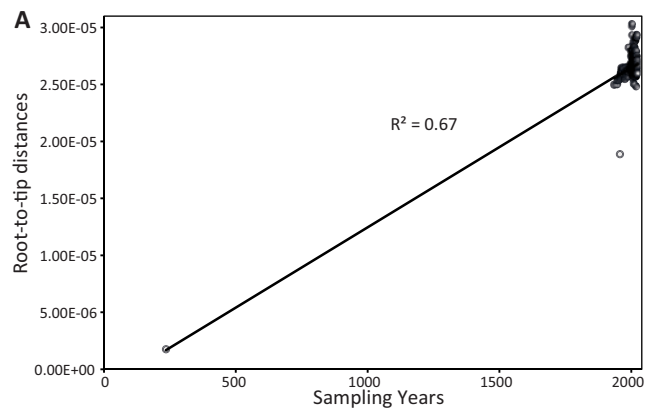Establishment and expansion of pig-enriched *Salmonella*
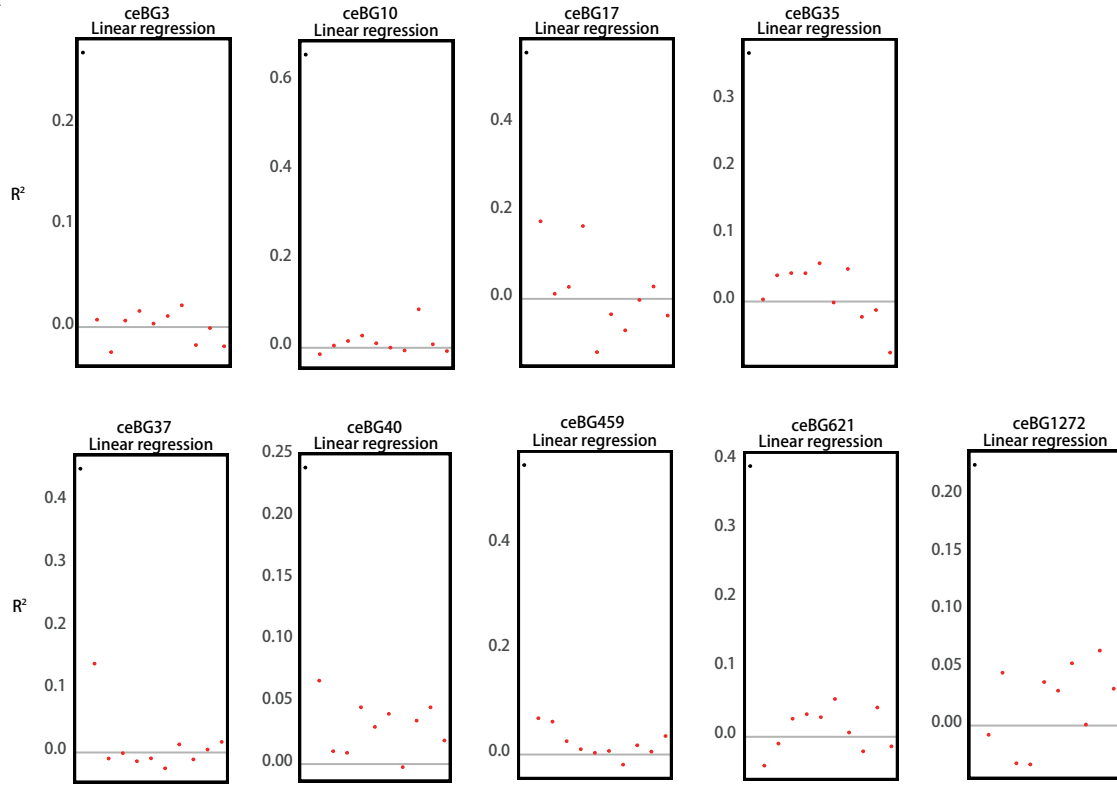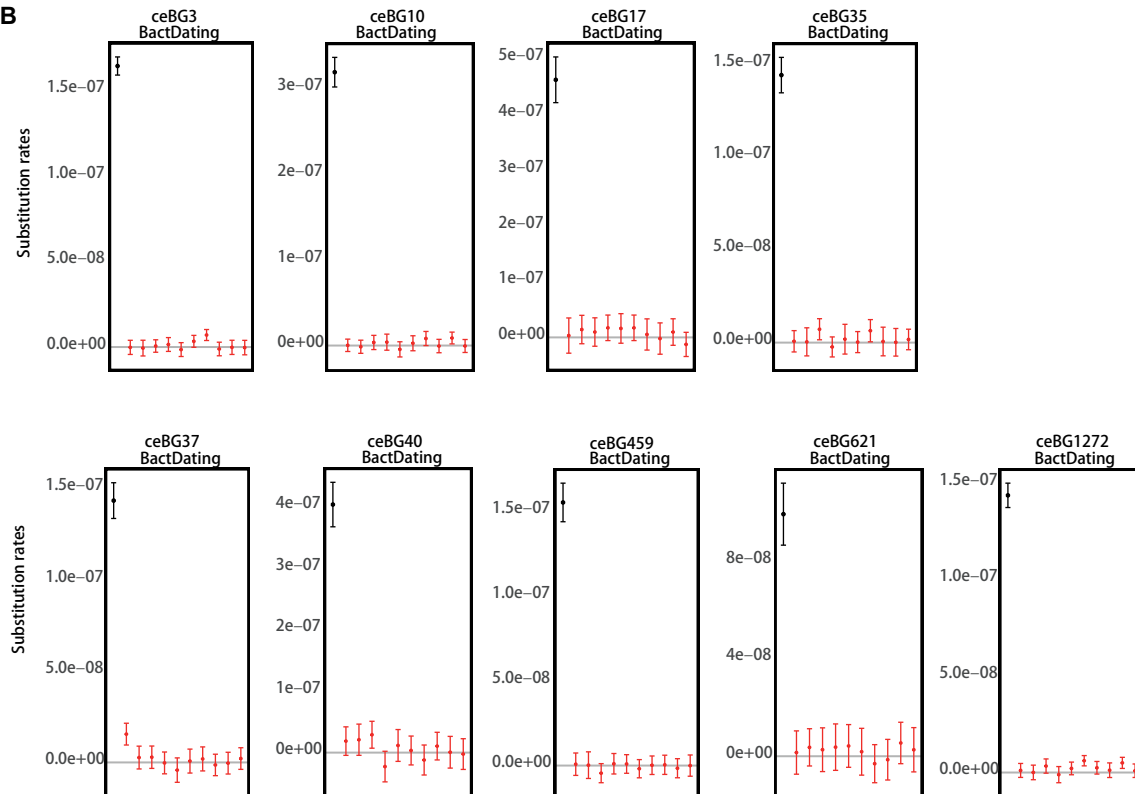
Antimicrobial Overuse

Increase of AMR level

Global Transportation

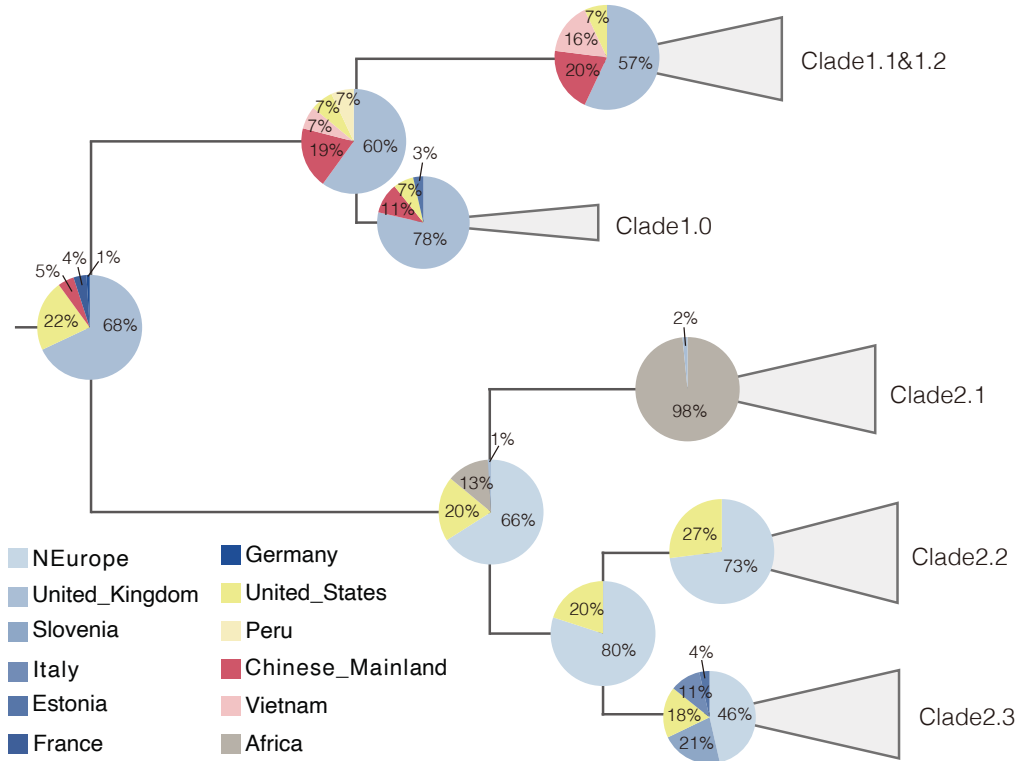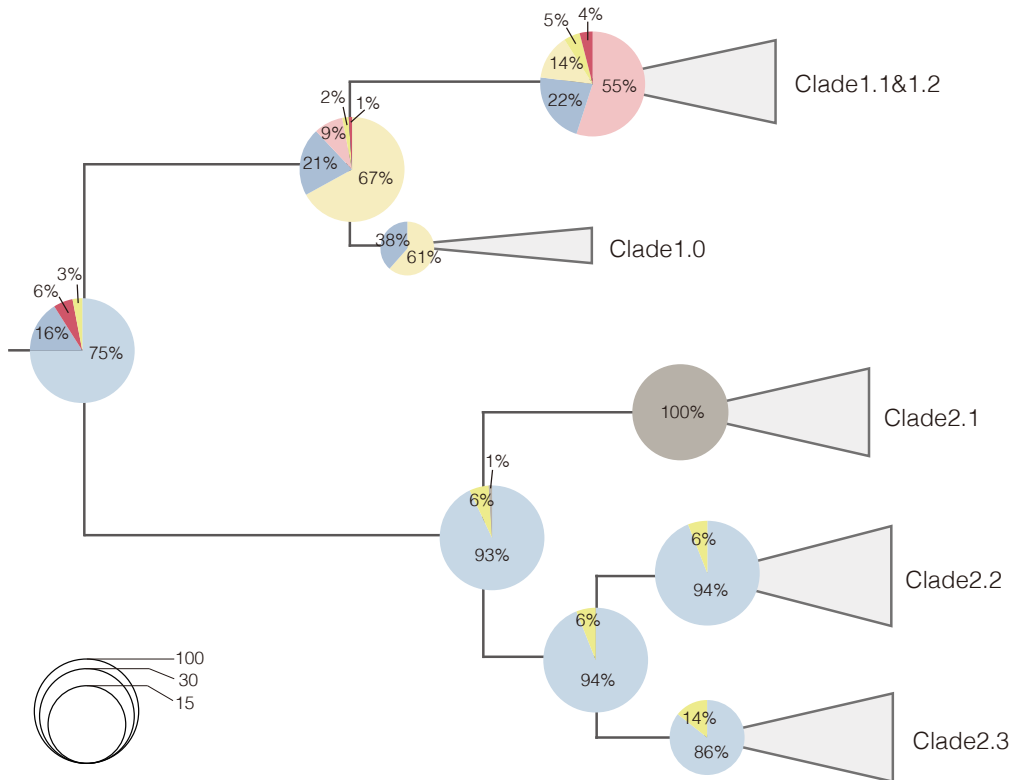Increase of international transmissions

Clade

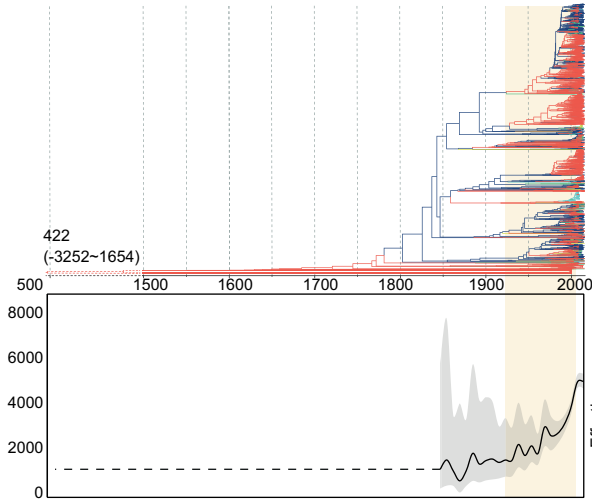| 1.0 | 2.2 |
1.1	2.2.1
1.1.1	2.2.2
1.1.2	2.2.3
1.2	2.2.4
1.2.1	2.3
1.2.2	2.3.1
1.2.3	2.3.2
2.0	2.3.3
2.1	2.3.4
2.1.1	2.3.5
2.1.2	2.3.6
2.1.3	2.3.7
2.1.4	3.1
	3.2

Region

East Asia
South Asia
Southeast Asia
West Europe
South Europe
North Europe
East Europe
South America
North America
Africa

Source

Human
Pig
Wild boar
Other animals
Environment

1e-5

A

Root-to-tip distances vs Sampling Years scatter plot with regression line, $R^2 = 0.67$

B Linear regression

R^2 plot

C BactDating

Substitution rates plot

A

ceBG3
Linear regression

ceBG10
Linear regression

ceBG17
Linear regression

ceBG35
Linear regression

ceBG37
Linear regression

ceBG40
Linear regression

ceBG459
Linear regression

ceBG621
Linear regression

ceBG1272
Linear regression

R^2

B

ceBG3
BactDating

ceBG10
BactDating

ceBG17
BactDating

ceBG35
BactDating

ceBG37
BactDating

ceBG40
BactDating

ceBG459
BactDating

ceBG621
BactDating

ceBG1272
BactDating

Substitution rates

A **Random sampling of 10 isolates from each country / region**

Clade1.1&1.2

Clade1.0

Clade2.1

Clade2.2

Clade2.3

NEurope
United_Kingdom
Slovenia
Italy
Estonia
France
Germany
United_States
Peru
Chinese_Mainland
Vietnam
Africa

B **Random sampling of 5 isolates from each country / region**

Clade1.1&1.2

Clade1.0

Clade2.1

Clade2.2

Clade2.3

A ceBG3(Derby)

422
(-3252~1654)

B ceBG1272(Choleraesuis)

-330
(-1948~220)

Effective population size

C ceBG17(Chailey)

1999-2014

1823
(1549~1934)

D ceBG35(Worthington)

1930~1953

1614
(997~1836)

E ceBG37(London)

1969~2001

1556
(916~1839)

F ceBG40(Cerro)

1986
(1979~1991)

2005~2012

G ceBG459(Johannesburg)

1937~1952

1627
(1587~1667)

H ceBG621(Ohio)

1959~1975

1690
(1152~1865)

I ceBG10(Adelaide)

1984~2007

1713
(1678~1747)

Human
Avian
Food/Environment
Bovine
Other Animals
Pig

A ceBG10(Adelaide)
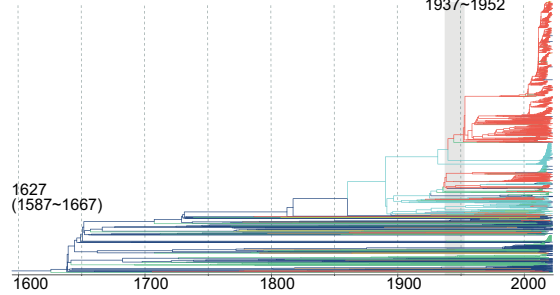B ceBG1272(Choleraesuis)
C ceBG17(Chailey)
D ceBG3(Derby)
E ceBG35(Worthington)
F ceBG37(London)
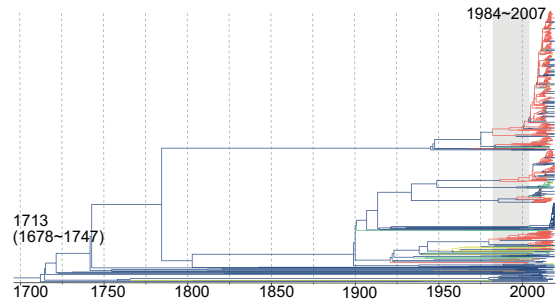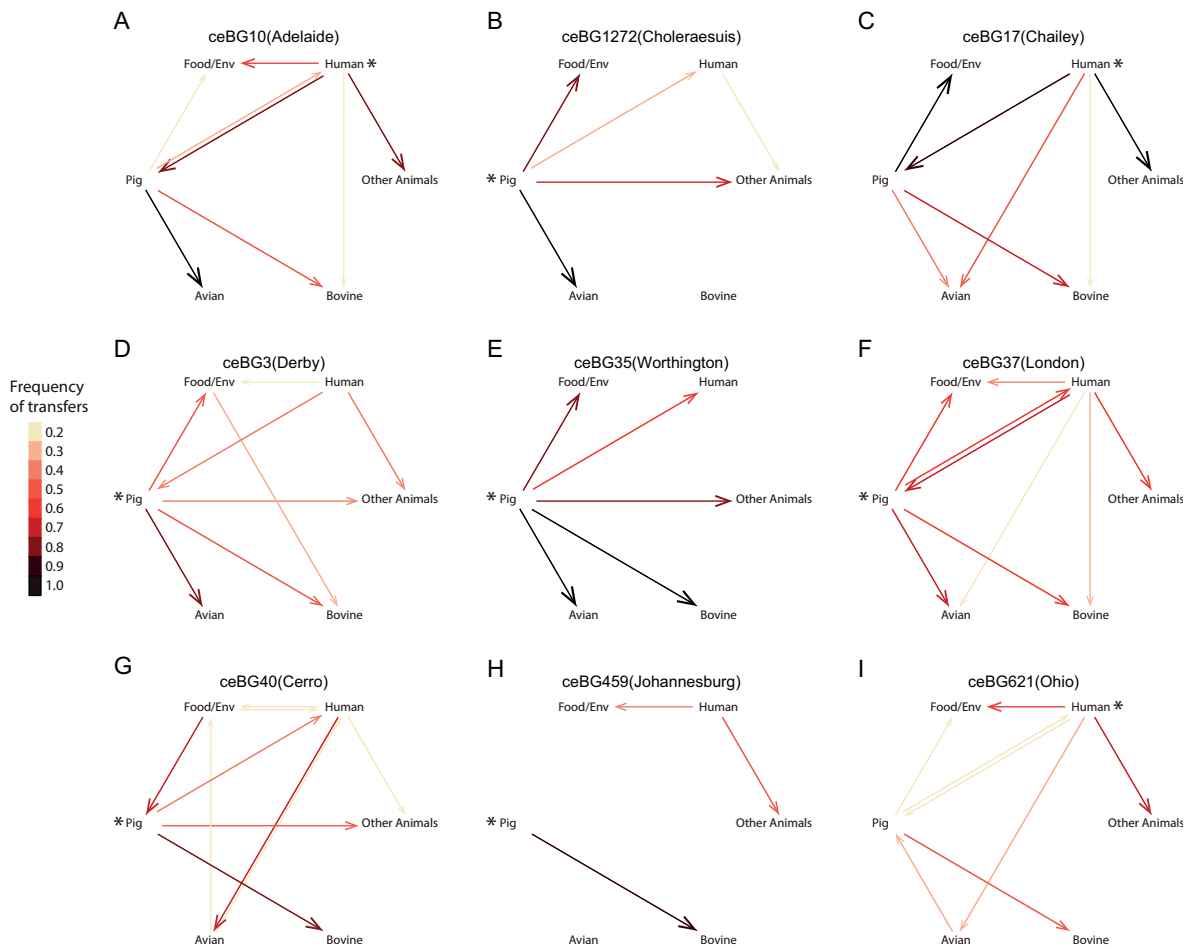G ceBG40(Cerro)
H ceBG459(Johannesburg)
I ceBG621(Ohio)

Frequency of transfers
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

A — Pork offal, frozen — R=0.94
B — Pork, frozen — R=0.96
C — Pork, frozen cut — R=0.93
D — Pork, preserved — R=0.89
E — Pork offal, fresh — R=0.90
F — Pig fat — R=0.89
G — Pork, fresh cuted — R=0.72
H — Pork, fresh — R=0.70
I — Pig carcasses — R=0.56
J — Pig, live breeding — R=0.53
K — Pig, live, less 50kg — R=0.52
L — Pig, live, more 50kg — R=0.46