

RESEARCH ARTICLE

Open Access



# Public health implications of *Yersinia enterocolitica* investigation: an ecological modeling and molecular epidemiology study

Yuan Yue<sup>1,2,3†</sup> , Jinxin Zheng<sup>4,5†</sup>, Mei Sheng<sup>6</sup>, Xiang Liu<sup>6</sup>, Qiong Hao<sup>6</sup>, Shunxian Zhang<sup>5,7</sup>, Shuai Xu<sup>2</sup>, Zhiguo Liu<sup>2</sup>, Xuexin Hou<sup>2</sup>, Huaiqi Jing<sup>2</sup>, Yang Liu<sup>8</sup>, Xuezhong Zhou<sup>1\*</sup> and Zhenjun Li<sup>2\*</sup>

## Abstract

**Background** *Yersinia enterocolitica* has been sporadically recovered from animals, foods, and human clinical samples in various regions of Ningxia, China. However, the ecological and molecular characteristics of *Y. enterocolitica*, as well as public health concerns about infection in the Ningxia Hui Autonomous Region, remain unclear. This study aims to analyze the ecological and molecular epidemiological characteristics of *Y. enterocolitica* in order to inform the public health intervention strategies for the contains of related diseases.

**Methods** A total of 270 samples were collected for isolation [animals ( $n = 208$ ), food ( $n = 49$ ), and patients ( $n = 13$ )], then suspect colonies were isolated and identified by the API20E biochemical identification system, serological tests, biotyping tests, and 16S rRNA-PCR. Then, we used an ecological epidemiological approach combined with machine learning algorithms (general linear model, random forest model, and eXtreme Gradient Boosting) to explore the associations between ecological factors and the pathogenicity of *Y. enterocolitica*. Furthermore, average nucleotide identity (ANI) estimation, single nucleotide polymorphism (SNP), and core gene multilocus sequence typing (cgMLST) were applied to characterize the molecular profile of isolates based on whole genome sequencing. The statistical test used single-factor analysis, Chi-square tests, *t*-tests/ANOVA-tests, Wilcoxon rank-sum tests, and Kruskal–Wallis tests.

**Results** A total of 270 isolates of *Yersinia* were identified from poultry and livestock ( $n = 191$ ), food ( $n = 49$ ), diarrhoea patients ( $n = 13$ ), rats ( $n = 15$ ), and hamsters ( $n = 2$ ). The detection rates of samples from different hosts were statistically different ( $\chi^2 = 22.636$ ,  $P < 0.001$ ). According to the relatedness clustering results, 270 isolates were divided into 12 species, and *Y. enterocolitica* ( $n = 187$ ) is a predominated species. Pathogenic isolates made up 52.4% (98/187), while non-pathogenic isolates made up 47.6% (89/187). Temperature and precipitation were strongly associated with the pathogenicity of the isolates ( $P < 0.001$ ). The random forest (RF) prediction model showed the best performance. The prediction result shows a high risk of pathogenicity *Y. enterocolitica* was located in the northern, northwestern, and southern of the Ningxia Hui Autonomous Region. The *Y. enterocolitica* isolates were classified into 54 sequence types (STs) and 125 cgMLST types (CTs), with 4/O:3 being the dominant bioserotype in Ningxia. The dominant STs and dominant CTs of pathogenic isolates in Ningxia were ST429 and HC100\_2571, respectively.

<sup>†</sup>Yuan Yue and Jinxin Zheng equally contributed as the first authors

\*Correspondence:

Xuezhong Zhou  
zhouxuezhong@nxu.edu.cn  
Zhenjun Li  
lizhenjun@icdc.cn

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



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Conclusions** The data indicated geographical variations in the distribution of STs and CTs of *Y. enterocolitica* isolates in Ningxia. Our work offered the first evidence that the pathogenicity of isolates was directly related to fluctuations in temperature and precipitation of the environment. CgMLST typing strategies showed that the isolates were transmitted to the population via pigs and food. Therefore, strengthening health surveillance on pig farms in high-risk areas and focusing on testing food of pig origin are optional strategies to prevent disease outbreaks.

**Keywords** *Yersinia enterocolitica*, Ecological, Machine learning, Molecular epidemiology, Core genome multilocus sequence typing, Ningxia, China

## Background

*Yersinia enterocolitica* is a common zoonotic pathogen that is widely found in soil, water, animals, and various foods [1, 2]. *Y. enterocolitica* is heterogeneous and consists of six biotypes: 1A, 1B, 2, 3, 4, and 5. Biotype 1B is regarded as highly pathogenic, and biotype 1A is regarded as being non-pathogenic in humans; other biotypes are considered to have low pathogenicity [3]. Enteric yersiniosis has been reported globally, with the most severe epidemic in Europe, where it was the fourth most common zoonotic disease in 2019 [4]. New Zealand has a high rate of yersiniosis notification, which is on the rise [5]. There were two outbreaks of *Y. enterocolitica* in China in the mid-1980s, resulting in over 500 infections [6]. Surveillance of the disease in China ceased for nearly 20 years, leading clinicians to rarely consider yersiniosis in the diagnosis of gastrointestinal diseases, and testers rarely being able to provide a basis for accurate diagnosis [7]. *Y. enterocolitica* was included in the national food contaminant surveillance network until 2016 [8]. The lack of mandatory government oversight in many countries may lead to an underestimation of the incidence of gastrointestinal disorders caused by *Y. enterocolitica* [9, 10]. Such cases reflect a general lack of attention to the disease at a domestic and international level, which can lead to untimely treatment and misdiagnosis, resulting in chronic and prolonged infection. As such, research on *Y. enterocolitica* is essential and urgent.

The emergence and re-emergence of zoonotic diseases have been recognized as being driven by variability at the human-animal-environment interface [11]. However, this comprehensive investigation is still lacking in almost all zoonotic systems, as we lack empirically validated models of ecological interactions between humans, animal reservoirs, and key environmental drivers [12]. In order to clarify the complexity of zoonotic spills, field monitoring and modelling methods that link human-animal contact with disease dynamics within animal reservoirs are essential [13]. The surveillance survey results of diseases caused by *Y. enterocolitica* in China through consecutive years have indicated that its hosts include livestock, poultry, and rodents [14]. Pigs are an essential source of *Y. enterocolitica* infections [15]. *Y. enterocolitica* is

transmitted via the faecal-oral route as well as through contact with animals (including farm animals, domestic pets, and wild animals) and contaminated food [14, 16]. Previous studies have shown that the separation rate of domestic animals (pigs and dogs) was significantly and positively correlated with altitude, mean temperature, and precipitation, significantly and negatively correlated with altitude and mean temperature, and significantly and positively correlated with rainfall [14]. However, the integrated picture of host-environment interactions and the resulting spread and spillover of *Y. enterocolitica* is far from clear. Yersiniosis is an example of a health threat that affects humans, animals, and our shared environment, and controlling outbreaks of yersiniosis highlights the importance of a sustainable One Health approach. A crucial component of the One Health approach is identifying and monitoring the prevalence of zoonotic diseases in an area. However, information about the epidemiology of *Y. enterocolitica* in Ningxia is scarce. In this area, *Y. enterocolitica* has been sporadically recovered from animals, foods, and human clinical samples.

The primary objective of this study was to explore the ecological factors and molecular characteristics that influence the spread and pathogenicity of *Y. enterocolitica*. An exogenous association between ecological factors and the pathogenicity of isolates in Ningxia was demonstrated using a predictive modelling approach. The distribution and molecular characteristics within *Y. enterocolitica* populations in the region were explored through correlated genomic analyses. The public health implications of the host-human-environment relationship were explored in terms of the endogenous and exogenous causes of *Yersinia* transmission and provided a reference for the tracking and control of related zoonotic diseases. Health threats at the human-animal-environment interface are best addressed through an efficient and sustainable One Health approach.

## Methods

### Data sources

The 270 isolates in this experiment were obtained from five prefectures in Ningxia Hui Autonomous Region

(Yinchuan, Shizuishan, Wuzhong, Guyuan, and Zhongwei) from 2007 to 2019. The isolates were obtained from animals, food, and patients. Animal hosts included pigs, sheep, rats, cattle, chickens, and hamsters. Pharyngeal swabs, anal swabs, intestinal contents, faeces and food samples were collected. Patient samples were obtained from faeces of diarrhoea patients collected at monitoring sites. Collection of pharyngeal swabs, anal swabs, and intestinal contents from healthy sheep and fattening pigs in slaughters. Pharyngeal swabs, anal swabs, and faeces collected from animals on farms. Intestinal contents of rats and hamsters from hunted rats and hamsters. Foods were purchased at butchers, retail markets and supermarkets in the region. All samples were transported to the laboratory in separate sterile containers using a cooler. Isolates enrichment was performed in phosphate-buffered saline with sorbitol and bile salts (PSB) at 4 °C for 21 days.

Environmental and climatic data obtained from open-access data sources, including elevation, normalized difference vegetation index (NDVI), monthly mean temperature, and rainfall (Additional file 8: Table S1). The NDVI and elevation data were aligned over a grid with a spatial resolution of 5×5 km. Data from the survey samples were extracted from the corresponding geo-referenced locations. The data processing procedures described above were using software R 4.6.0 (Lucent Technologies, Jasmine Mountain, USA) with the ‘raster’ package.

#### Collection and identification of isolates

*Yersinia* was inoculated onto selective agar (CIN Agar; Oxoid, Basingstoke, UK/HKM, Guangzhou, PRC). A typical bulls-eye appearance (deep red centers surrounded by outer transparent zones) on CIN-selective agar plates was inoculated onto Kligler iron and urea media. We performed the identification of the isolates using the API20E biochemical identification system. The 16SrRNA fragment was amplified to confirm the species. The isolates were further distinguished by serotyping (*Y. enterocolitica* antisera set from the Institute of Chinese Biomedicine) [17] and biotyping (Bile Aesculin Agar; Oxoid, Basingstoke, UK. Brain Heart Infusion Agar; Oxoid, Basingstoke, UK. Tween 80; Amresco, USA. CaCl<sub>2</sub>; Sinopharm Chemical Reagent Co., Ltd, PRC. Biochemical Reaction Tablets; Rosco, Denmark) [19]. The biotype was determined by biochemical experiments on lipase activity, salicin, esculin hydrolysis, xylose, trehalose, indole production, ornithine decarboxylase, Voges-Proskauer test, pyrazinamidase activity, sorbose, inositol, and nitrate reduction [18, 19].

#### Whole-genome sequencing and assembly

Genomic DNA was extracted using a Wizard® Genomic DNA Purification Kit (Promega, United States) according to the manufacturer’s protocol. DNA concentration, quality, and integrity were determined using a 5400 Fragment Analyzer System (Agilent, United States) and a NanoDrop 2000 Spectrophotometer (Thermo, United States). All 270 isolates were sequenced using Illumina Nova technology. Raw reads were quality filtered and assembled. The conditions for filtering were as follows: (1) reads containing more than 40% low-quality bases (mass value ≤20) were removed; (2) reads with N beyond a certain proportion (the default was 10%) were removed; (3) reads that overlapped with the adapter sequence by more than 15 bp with fewer than three mismatches between them were removed. Assemblies were performed using SOAP denovo (<http://soapdenovo2.sourceforge.net/>) [20], SPAdes (<http://bioinf.spbau.ru/spades>), and Abyss (<https://www.bcgsc.ca/resources/software/abyss>) software. Assembly results were integrated using CISA software (<http://sb.nhri.org.tw/CISA/en/CISA>). The preliminary assembly results were complemented with gap-closing software to filter reads with low sequencing depths (less than 0.35 of the average depth) to remove homolane contamination, resulting in the final assembly results. The reference genomes of 26 species of the *Yersinia* genus were shown in Additional file 8: Table S2. The ANI was estimated for each pair of isolates using the 270 *Yersinia* genomes, the 26 reference genomes (see above), and fastANI v1.33 (<https://github.com/ParBLiSS/FastANI>) [21]. The reference genes for the *Yersinia* were the same as those listed in the previous section.

#### Phylogenetic and pan-genome analyses

*Y. enterocolitica* 8081 (GCA\_000009345.1) was used as the reference sequence, and the whole-genome sequences of 270 isolates were compared with the reference sequence using MUMmer 4.0 software (<http://mummer.sourceforge.net/>) [22]. All SNP loci of each strain were obtained. The following protocol was used: Integrate the position data of all SNP loci in the reference genome, filter SNPs from duplicated regions, and remove low-quality SNPs for subsequent analysis [23, 24]. Maximum likelihood (ML) trees were constructed using iQ-TREE (<http://www.iqtree.org/>) with 1000 bootstrap replicates. In addition, the whole genome sequences of 88 publicly available *Y. enterocolitica* isolates were downloaded from the NCBI database (<https://www.ncbi.nlm.nih.gov/>). These isolates were isolated from 11 countries: the United Kingdom ( $n=36$ ), New Zealand ( $n=19$ ), France ( $n=10$ ), Germany ( $n=8$ ), the United States ( $n=3$ ), Italy ( $n=4$ ),

Ireland ( $n=2$ ), China ( $n=2$ ), Spain ( $n=1$ ), Greece ( $n=1$ ), and Australia ( $n=1$ ); and 1 unspecified as shown in Additional file 8: Table S3. An ML tree of *Y. enterocolitica* (this study and NCBI database) was constructed as the reference genome *Y. enterocolitica* 8081 (GCA\_000009345.1) using roary v3.13.0 software (<https://github.com/sanger-pathogens/Roary>). The trees were then visualized and annotated in iTOL (<https://itol.embl.de/>) (using their associated scripts that were obtained after processing the previous series) [24].

### Virulence profiles

The assembled genomes of the isolates in this study were inserted into the Virulence Factor Database (VFDB; MOH Key Laboratory of Systems Biology of Pathogen, Institute of Pathogen Biology, Beijing, <http://www.mgc.ac.cn>) with default parameter selection.

### ST and core genome MLST determination

Sequence types (STs) and cgMLST (CTs) of *Y. enterocolitica* were assigned using EnteroBase (<http://enterobase.warwick.ac.uk/species/index/yersinia>) [25]. The McNally 7 Gene scheme contained *aarE*, *dfp*, *galR*, *glnS*, *hemA*, *rfaE*, and *speA* [26]. CgMLST scheme, cgMLST V1+HeirCC V1 (Hierarchical cgMLST clustering) was based on the profiles of 1553 coding loci. The minimum spanning tree (MST) and neighbor-joining (NJ) trees were constructed using isolates from the Ningxia region and those publicly available in the NCBI database with GrapeTree v2.1 (<https://github.com/achtman-lab/GrapeTree>).

### Statistical analysis

Single-factor analysis was conducted to compare the differences of ecological environment factors between the host, biotype, serotype, and pathogenicity of *Yersinia*. Chi-square tests were used for categorical variables, while *t*-tests or ANOVA-tests were applied for normal distribution of continuous variables. If there exists non-normal distribution, Wilcoxon rank-sum tests and Kruskal–Wallis tests were used. The significance level was set to 0.05. The study also using the ecological environment factors to predict the risk of pathogenicity of *Yersinia*. We have split our data into training dataset (80%) and testing dataset (20%), and utilized general linear model (GLM), random forest model (RF), and XGBOOST (eXtreme Gradient Boosting) models to evaluate methodological improvements (Additional file 8: Table S4). The area under the receiver operating characteristic (ROC) curve was used to evaluate the model performance. The model was simulated by ‘caret’ package, and AUC (area under the curves) was performed by ‘PRROC’ package with statistical software R 4.6.0.

## Results

### Epidemic profile of *Yersinia* during 2007–2019

A total of 9031 samples were monitored from 2007 to 2019, with the detection rate of *Yersinia* ranging from 0.9% to 7.6% (Table 1). The highest detection rate was in 2014 (7.6%), eightfold higher than in 2013 (0.5%). The difference in positivity rates between years was statistically significant ( $\chi^2=40.282$ ,  $P<0.001$ ). No significant upward or downward trend in detection rates between years. Samples from different hosts were detected from 0.3% to 5.9%, with the highest detection rate of 5.9% for samples from poultry and livestock sources (Table 2). The detection rates of samples from different hosts were statistically different ( $\chi^2=22.636$ ,  $P<0.001$ ).

The isolates were obtained from five prefectures in Ningxia Hui Autonomous Region [Yinchuan ( $n=111$ ), Shizuishan ( $n=7$ ), Wuzhong ( $n=1$ ), Guyuan ( $n=3$ ), and Zhongwei ( $n=148$ )], from 2007 to 2019 (Fig. 1 and Additional file 8: Table S5). In total, 208 (77.0%) were of animal origin, 49 (18.2%) were of food origin, and 13 (4.8%) were of patient origin. Animal hosts included pig (150/208, 72.1%), sheep (32/208, 15.4%), rat (15/208, 7.2%), cattle (6/208, 2.9%), chicken (3/208, 1.4%), and hamster (2/208, 1.0%). The source of animal samples was mainly feces ( $n=102$ ), pharyngeal swabs ( $n=39$ ), anal swabs ( $n=7$ ), and intestinal contents ( $n=59$ ) of animals. Food was derived from meat products, comprising beef ( $n=25$ ), pork ( $n=10$ ), chicken ( $n=9$ ), lamb ( $n=3$ ), and fish ( $n=2$ ). Food ( $n=49$ ) comprised fresh meat ( $n=18$ ) and frozen meat ( $n=31$ ). Human samples were of fecal origin, with the majority ( $n=9$ ) coming from children (Fig. 1 and Additional file 8: Table S5). Using traditional phenotypic methods, 187/270 (69.3%) isolates were serotyped, with the most common serotypes being O:3 ( $n=84$ ), O:5

**Table 1** Detection rates of *Yersinia*, 2007–2019 in Ningxia Hui Autonomous Region, China

Year	Sample size	Positive samples	Positive rate (%)
2007	539	18	3.3
2008	118	2	1.7
2009	816	39	4.8
2010	660	7	1.1
2011	959	37	3.9
2012	1327	28	2.1
2013	217	2	0.9
2014	384	29	7.6
2015	578	14	2.4
2016	630	8	1.3
2017	852	20	2.4
2018	899	36	4.0
2019	1052	30	2.9



**Table 2** Detection rates of *Yersinia* in samples from different hosts, 2007–2019 in Ningxia Hui Autonomous Region, China

Host	Sample size	Positive samples	Positive rate (%)	Sample source
Livestock and poultry	3219	191	5.9	Feces, pharyngeal swabs, anal swabs, and intestinal contents
Food	511	49	3.7	Smear swabs
Diarrhoea patients	3824	13	0.3	Feces
Rat	1190	15	1.3	Intestinal contents
Hamster	287	2	0.7	Intestinal contents

( $n=52$ ), O:8 ( $n=24$ ), and O:9 ( $n=20$ ). In total, 83 isolates were reported as O: unidentifiable because the O-antigen reacted with more than one antiserum or with none of the antisera (Fig. 1).

#### Distribution of *Y. enterocolitica* by biotype, host, and serotype

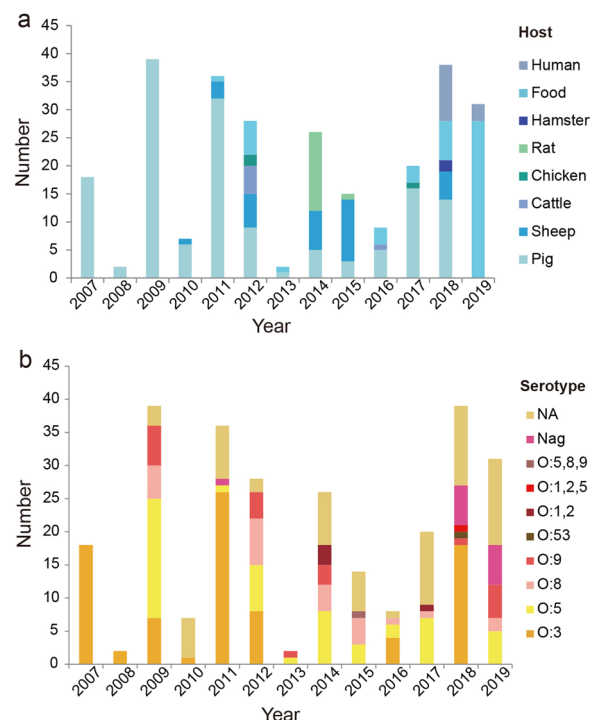
Of the 187 *Y. enterocolitica* isolates, 81.3% ( $n=152$ ) were of animal origin, the food source was 12.3% ( $n=23$ ) and the patient source was 6.4% ( $n=12$ ). Isolates of animal origin included 42.8% biotype 1A ( $n=65$ ), 50.0% biotype 4 ( $n=76$ ), 2.6% biotype 3 ( $n=4$ ) and 4.6% biotype 5 ( $n=7$ ). Of these, all isolates of biotype 5 were from sheep. In total, biotype 4 of isolates ( $n=84$ ) were from pig hosts. The dominant serotype of the isolates of animal origin was O:3. Food-derived isolates were 87.0% biotype 1A ( $n=20$ ) and 13.04% biotype 2 ( $n=3$ ). Patient-origin isolates included 33.3% biotype 1A ( $n=4$ ) and 66.7% biotype 4 ( $n=8$ ). There was no biotype 1B (Table 3 and Additional file 8: Table S6).

#### Phylogenetic analysis

Total of 270 *Yersinia* genomes were evaluated according to the standard 95–96% ANI [27]. Twelve species were delineated using a 95% ANI cut-off value: *Y. enterocolitica* (187/270, 69.3%), *Y. intermedia* (30/270, 11.1%), *Y. massiliensis* (30/270, 11.1%), *Y. mollaretii* (7/270, 2.6%), *Y. pekkanenii* (5/270, 1.9%), *Y. proxima* (4/270, 1.5%), *Y. alsatica* (2/270, 0.7%), *Y. frederiksenii* (1/270, 0.4%), *Y. kristensenii* (1/270, 0.4%), *Y. hibernica* (1/270, 0.4%), *Y. canariae* (1/270, 0.4%), and *Y. rochesterensis* (1/270, 0.4%) (Fig. 2). The clustering analyses was consistent with the results of the ANI analysis and identical separation into 12 distinct species as determined by BAPS (Bayesian analysis of population structure) (Fig. 2).

The relatedness clustering presented by the *Y. enterocolitica* ML tree showed a direct relationship with the biotype. Biotypes 1A, 1B, and 5 isolates formed discrete clusters, whereas biotypes 2, 3, and 4 isolates consisted of closely related but distinct lineages, confirmed by BAPS

clustering. Pathogenic isolates from Ningxia were genetically more distant from the reference genome than isolates from other countries and regions (no biotype 1B isolates in Ningxia). Isolate LC20 from Zhejiang had the longest genetic distance with the reference genome compared to the other isolates in China. Biotype 1A isolates exhibited no geographical differences and a broader range of host species and serotypes. The hosts of biotypes 4 and 5 isolates in Ningxia were pigs and sheep, respectively, whereas the hosts in the other regions were humans and hares. Biotype 3 isolates were O:3 in Ningxia, compared to O:5,27 and O:6,30 in the other countries (Fig. 2).



**Fig. 1** Number of *Yersinia* spp. isolates isolated from the Ningxia Hui Autonomous Region between 2007 and 2019. **a** Host distribution of isolates. **b** Serotype distribution of isolates. NA, not applicable; Nag, nonagglutinative

**Table 3** Serotype and biotype distribution of *Y. enterocolitica* isolates

		Animal (n = 152)					Food (n = 23)	Human (n = 12)
		Pig (n = 121)	Sheep (n = 17)	Cattle (n = 6)	Rat (n = 5)	Chicken (n = 3)		
Biotype	1A	41/33.9%	10/58.8%	6/100%	5/100%	3/100%	20/87.0%	4/33.3%
	2	0	0	0	0	0	3/13.0%	0
	3	4/3.3%	0	0	0	0	0	0
	4	76/62.8%	0	0	0	0	0	8/66.7%
	5	0	7/41.2%	0	0	0	0	0
Serotype	O:3	78/64.5%	0	0	0	0	0	5/41.7%
	O:5	30/24.8%	2/11.8%	6/100%	5/100%	2/66.7%	6/26.1%	0
	O:8	5/4.1%	12/70.6%	0	0	1/33.3%	5/21.7%	1/8.3%
	O:9	8/6.6%	2/11.8%	0	0	0	8/34.8%	1/8.3%
	O:53	0	0	0	0	0	1/4.4%	0
	O:1,2,5	0	0	0	0	0	0	0
	O:5,8,9	0	1/5.9%	0	0	0	0	1/8.3%
	Nag	0	0	0	0	0	3/13.0%	4/33.3%

The data indicate the number of isolates and the percentage of them. *Nag* nonagglutinative

### Associations between ecological factors and pathogenicity

Temperature, precipitation, altitude, and NDVI were highly statistically significant to pathogenicity ( $P < 0.001$ ) (Fig. 3a–d). The ambient temperature of the collection locations for pathogenic isolates was lower than that of the non-pathogenic isolates ( $12.70 \pm 6.39$  °C vs.  $16.85 \pm 6.05$  °C); the precipitation of the collection locations for the pathogenic isolates was lower than that of the non-pathogenic isolates ( $1.34 \pm 1.29$  mm vs.  $2.54 \pm 1.42$  mm). The median elevation of the collection locations for pathogenic and non-pathogenic isolates was 1111.00 m and 1816.00 m, respectively; the median NDVI was 0.35 and 0.19, respectively. Also, except for the host of the isolates in precipitation factors, the ecological factors and pathogenicity were statistically highly significant with biotype, serotype of the isolates, and host of the isolates (Fig. 3, Additional file 1: Fig. S1, Additional file 2: Fig. S2, Additional file 3: Fig. S3).

Then, we used three models to predict the pathogenic of *Y. enterocolitica* with ecological environment factors. The AUC of the training and test sets of all three models exceeded 0.9, particularly the AUC of RF, and XGBOOST, indicating good evaluation performance (Fig. 3e, f). The contributions of ecological factors to the risk of *Y. enterocolitica* pathogenicity were shown in Additional file 4: Fig. S4, which indicated the prediction

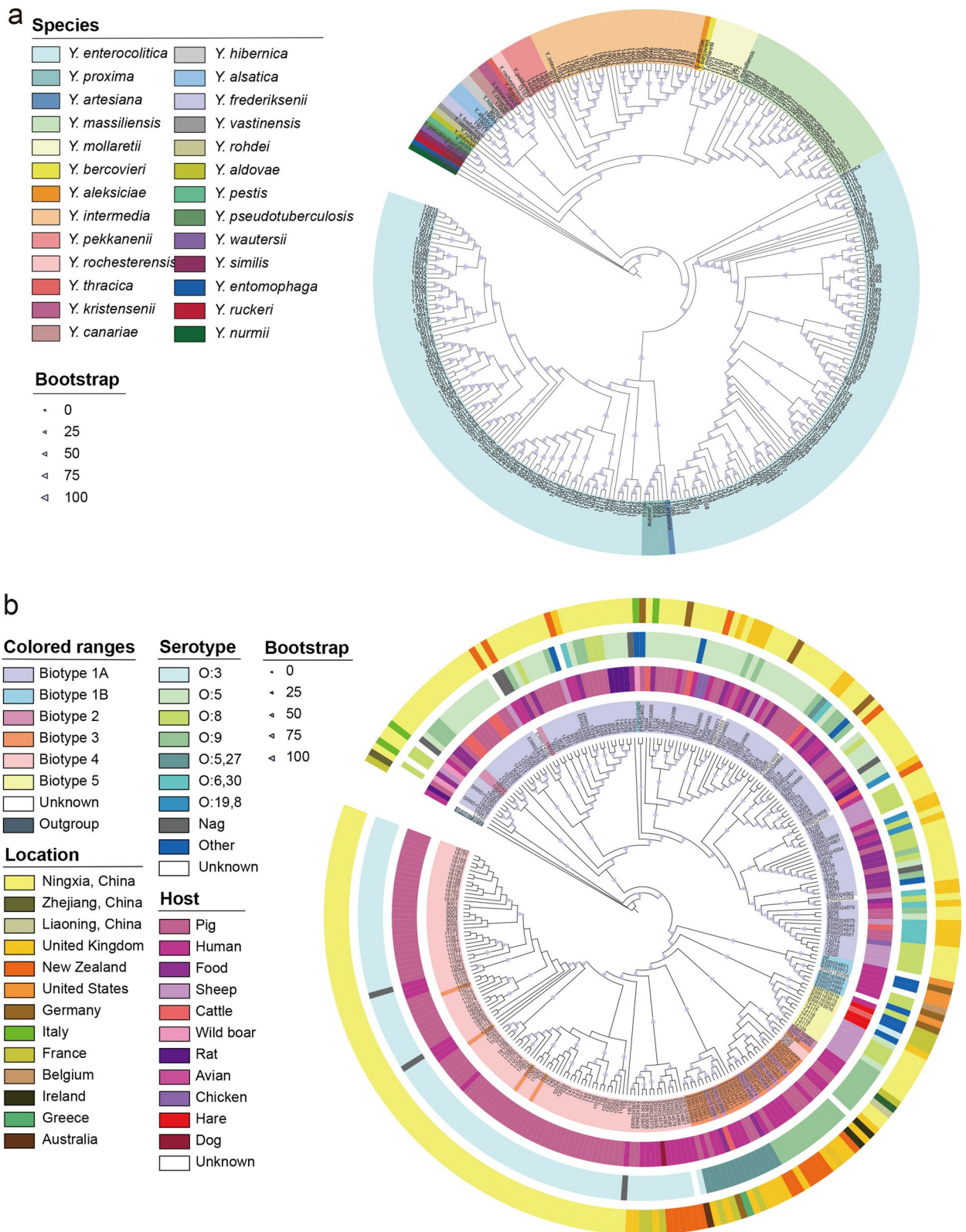
of *Y. enterocolitica* pathogenicity was the combination of all ecological factors. Finally, we used those performed models to evaluate the risk of *Y. enterocolitica* pathogenicity among our study regions (Additional file 5: Fig. S5). The results indicated that the highest risk area was concentrated in Yinchuan City and Shizuishan City in northern Ningxia, Zhongning City, and Zhongwei County in central Ningxia, and the southeastern part of Guyuan City and Haiyuan County in southern Ningxia. The predicted risk also varied between sources, with animals and humans suggesting a wider range of risk areas.

### Virulence profiles

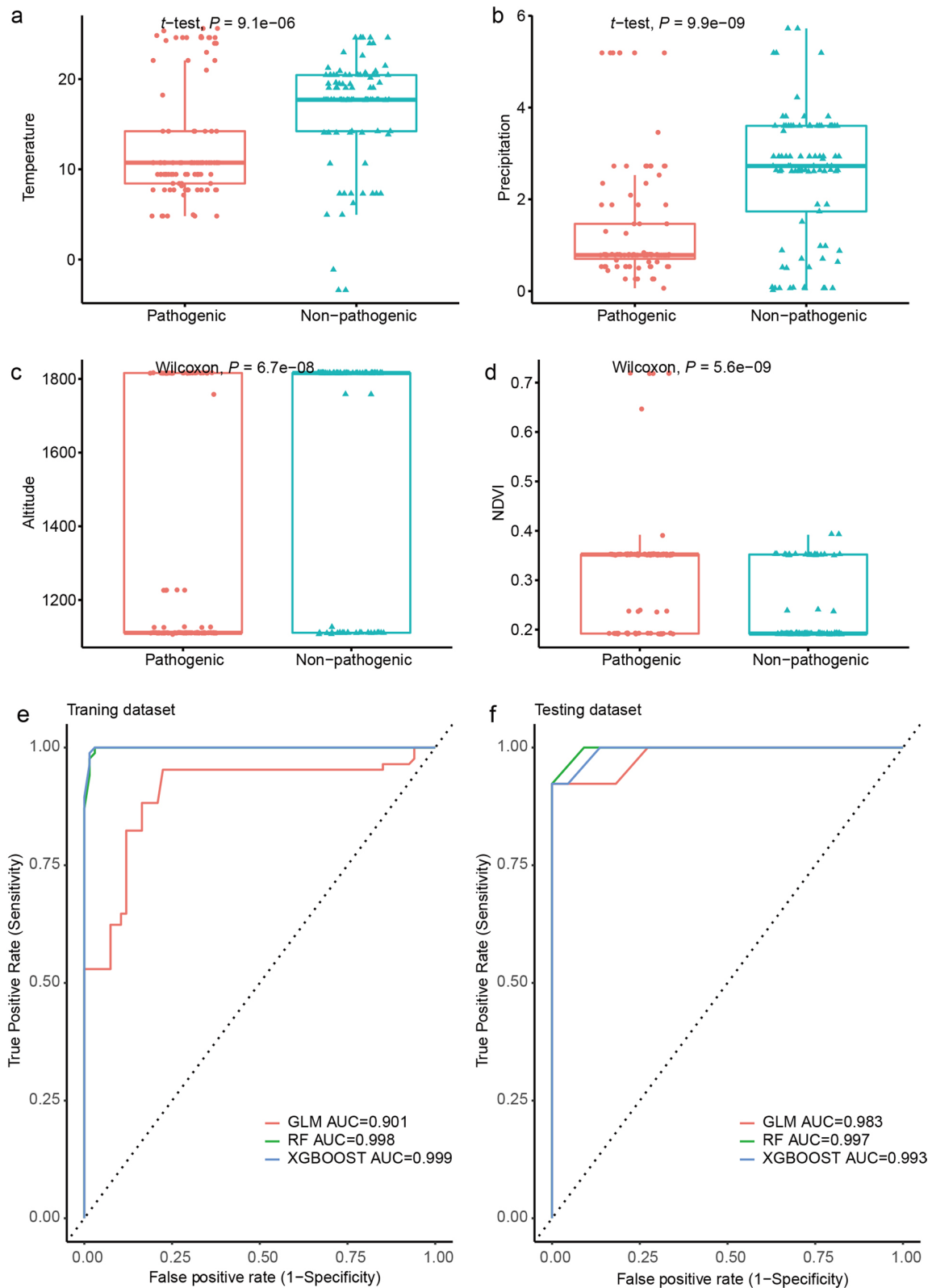
*Y. enterocolitica* isolates were annotated with 130 virulence genes in 5 categories: flagella, invasion, O-antigen, Yersiniabactin, and T3SS (Fig. 4). There were 53 flagella genes, most of which were *che*, *flh*, *flg* and *fli* genes. The O-antigen was represented by 27 genes, mostly *mrk* and *wbc* genes. The Yersiniabactin was made up of 9 genes, the majority of which were *ybt* gene. T3SS included 41 genes. The major ones were *ysc*, *yop*, *icr*, and *sys* genes. The distribution of virulence factors was closely related to the biotype. Most biotype 1A isolates lacked the virulence factors of invasion, O-antigen, Yersiniabactin, and T3SS. In contrast to pathogenic isolates of other biotypes, biotype 1B isolates were presented in Yersiniabactin.

(See figure on next page.)

**Fig. 2** SNP-based maximum likelihood (ML) trees of *Yersinia* spp. isolates. **a** The SNP-based ML tree was built from a recombination-filtered alignment of the whole genome SNP (wgSNP) present in 270 isolates. The ML tree was built using the GTR + F + ASC + G4 model, with 1000 bootstraps based on 1,563,073 SNPs, with *Y. enterocolitica* 8081 (GCA\_000009345.1) was used as the reference sequence. **b** The cgSNP-based ML tree was built from 187 *Y. enterocolitica* isolates of this study and 88 public data. *Y. enterocolitica* 8081 (GCA\_000009345.1) was used as the reference sequence. *Y. pseudotuberculosis* (GCA\_900637475.1) was used as outgroups

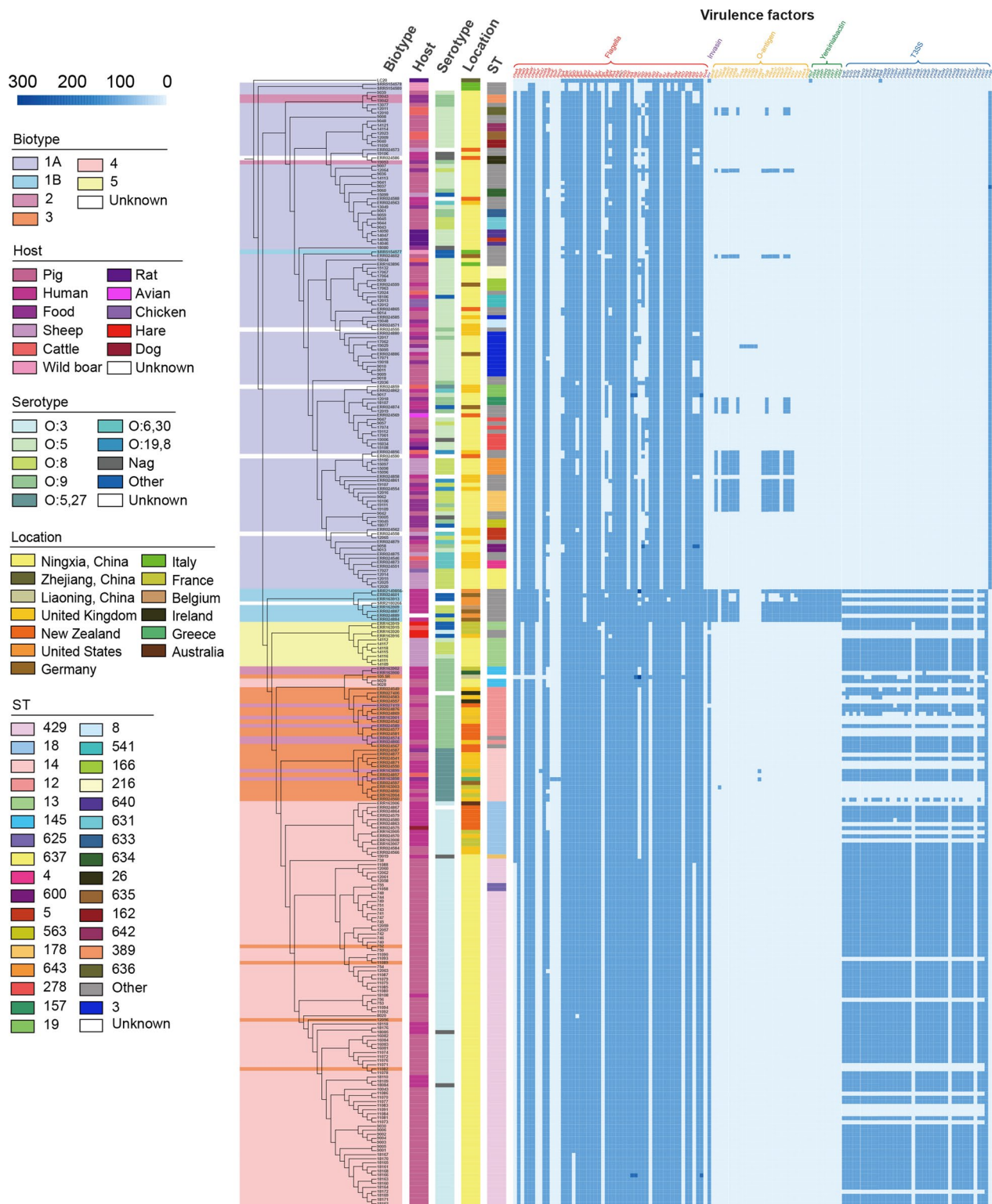


**Fig. 2** (See legend on previous page.)



**Fig. 3** Differences in pathogenicity of *Yersinia enterocolitica* isolates and ecological factors and the development of predictive models. **a** Temperature. **b** Precipitation. **c** Altitude. **d** NDVI. **e** Training dataset. **f** Testing dataset





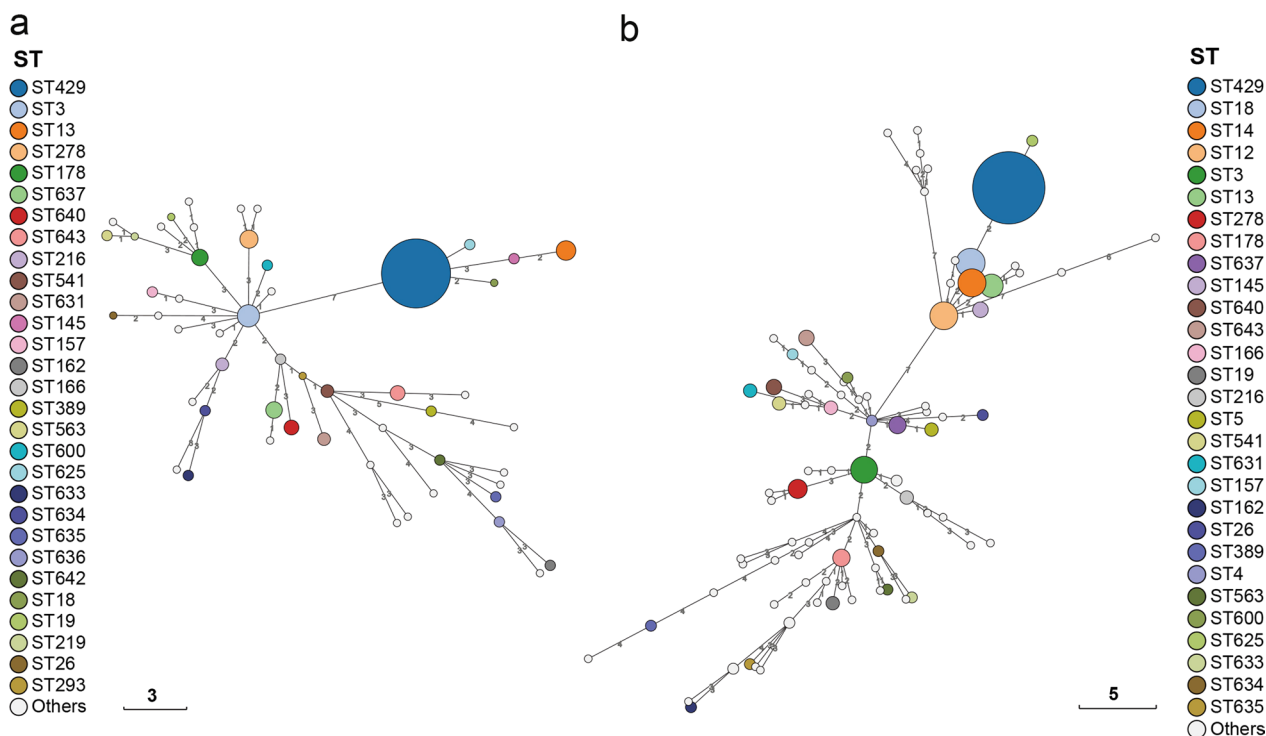
**Fig. 4** Distribution of virulence genes in *Yersinia enterocolitica* isolates. The heat map on the right depicted the host, serotype, STs, location, and virulence genes of the isolates

### ST and cgMLST typing

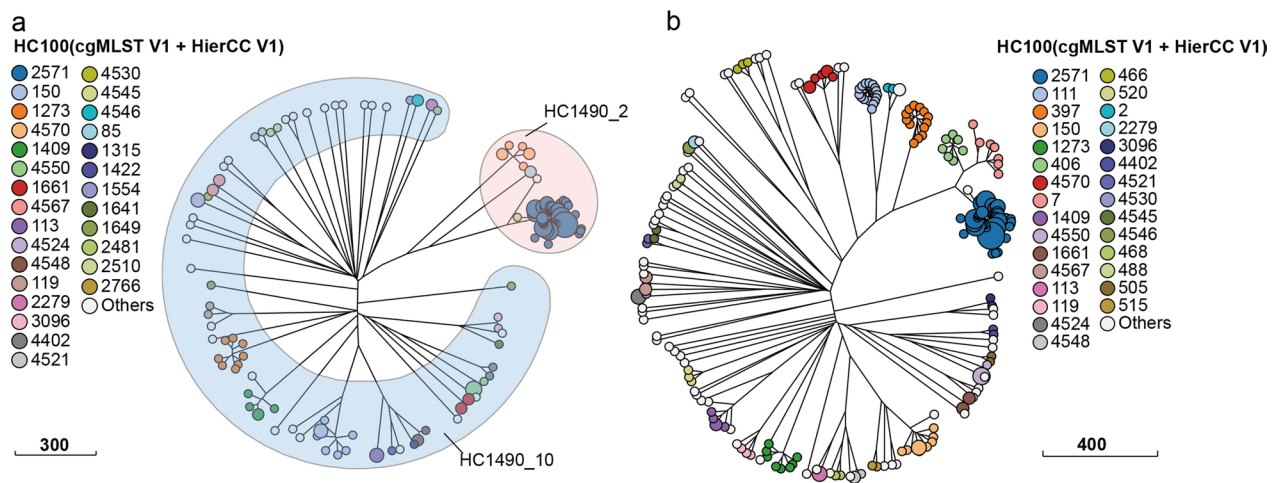
In 187 *Y. enterocolitica* isolates, 54 STs were detected using multilocus sequence typing. ST429 was the most common, accounting for 42.3% of all STs (79/187) (Fig. 5 and Additional file 8: Table S6). ST429 isolates included biotype 4 ( $n=79$ ) and biotype 3 ( $n=2$ ), and they were closely related to serotype O:3. ST3 ( $n=9$ ), ST278 ( $n=6$ ), ST178 ( $n=5$ ), ST637 ( $n=5$ ), ST640 ( $n=4$ ), ST643 ( $n=4$ ), and ST216 ( $n=3$ ) were biotype 1A isolates with serotypes O:5, O:8, O:9. ST13 isolates ( $n=7$ ) were biotype 5, with serotypes O:5, O:8, and O:9. ST3, ST12, ST14, ST18, and ST13 were the principal STs of foreign isolates, which differed from those in the Ningxia region (Additional file 8: Tables S3, S6, and Fig. 5). The biotypes and serotypes corresponding to ST12 and ST4 were bioserotypes 1A/O:6,30 and 2,3,4/O:9 (Additional file 6: Fig. S6).

An NJ tree and an MST based on cgMLST analysis of 187 *Y. enterocolitica* isolates were constructed. These 1,553 cgMLST target genes were randomly distributed across the genome. *Y. enterocolitica* isolates were divided into 125 cgMLST types (CTs) (Additional file 8: Table S6 and Fig. 6). CgMLST analysis revealed the core genome diversity of isolates with the same ST from 0 to 84 allelic differences. The NJ tree of 187 *Y. enterocolitica* isolates from Ningxia indicated the names of two microclades

of the HC1490 (Hierarchical clustering) cluster (Fig. 6). HC1490\_10 and HC1490\_2 were the primary phylogenetic branches. These two microclades of the HC1490 cluster were consistent with the results of the *Y. enterocolitica* ML tree. HC1490\_2 was strongly associated with biotypes 3, 4, and 5 isolates. HC1490\_10 was closely related to biotype 1A isolates. The 125 CTs present in the 187 isolates from the Ningxia region clustered to form 54 microclades of HC100. Of these, HC100\_2571 was the principal microclades, and the isolates were all from this study. The NJ tree constructed from isolates from the Ningxia region and those publicly available in the database showed that several microclades of the HC100 cluster were significantly associated with serotypes, hosts, and countries. HC100\_2571 isolates were all obtained from this study and were of bioserotype 4/O:3 and 3/O:3, hosts were pigs and humans. HC100\_406 isolates were 4/O:3, mostly from patients in New Zealand. The serotype of HC100\_397 isolates was O:5,27, with hosts of pigs, humans, and food. HC100\_2 and HC100\_111 isolates were 3/O:9 and 4/O:9, respectively, mainly from pigs and humans in the UK and New Zealand. HC100\_4570 isolates were biotype 5, obtained from sheep in Ningxia. HC100\_466 isolates were biotype 1B, which were isolated from patients. HC100\_150 isolates from pigs and patients



**Fig. 5** The minimum spanning trees (MST) of *Yersinia enterocolitica* isolates. The circle size was proportional to the number of isolates. Links between circles were represented according to the number of allelic differences between STs. **a** Isolates from this study. **b** Isolates from this study and the public database



**Fig. 6** The neighbor-joining (NJ) tree of *Yersinia enterocolitica* isolates based on cgMLST. The circle size was proportional to the number of isolates. Clusters generated using the hierarchical clustering method from EnteroBase and using a 100 cgMLST allele distance (HC100) were represented by circles. **a** Isolates from this study. **b** Isolates from this study and the public database

in Ningxia were 1A/O:5. The isolates HC100\_1273 were 1A/O:5 and obtained from pigs, cattle, and poultry in the Ningxia region. The mostly hosts of the biotype 1A isolates were pigs, cattle, sheep, rats, poultry, and food. The most common serotypes were O:5, O:8, O:9, and O:6,30 (Fig. 6 and Additional file 7: Fig. S7).

## Discussion

Ningxia Hui Autonomous Region is located in the upper reaches of the Yellow River in western China, where the largest Hui nationality in the country resides. Of the five cities in Ningxia, Guyuan and Wuzhong are Muslim-concentrated areas where livestock farms are not allowed to raise pigs, resulting in a small sample size in these two areas. Yinchuan and Zhongwei are areas with scattered Muslim populations, and 95.9% of the samples (259/270) in this study were collected from these two regions.

*Yersinia* presented a diversified epidemiological trend in Ningxia over a 12-year period, with a pattern of the diminished predominance of prevalent isolates and the coexistence of multiple subtypes. There is no significant upward or downward trend in detection rates between years. The detection rates of samples from different hosts were statistically different ( $\chi^2 = 22.636$ ,  $P < 0.001$ ), as well as 72.1% of poultry and livestock samples originated from pigs, implying that pigs were the primary source of *Yersinia* in Ningxia. *Y. enterocolitica* is present in the oral cavity (especially the tonsils and submaxillary lymph nodes), intestine, and feces of pigs [28, 29]. *Y. enterocolitica* was easily transmitted between pigs via the faecal-oral route and may also be spread through direct contact between humans and animals or the environment (water and soil) [30–32]. Therefore, people involved in swine farming and

processing or those with livestock farms in the vicinity of their living area are at greater risk of *Y. enterocolitica* infection. Food (18.2%) was another major source of *Yersinia* infection in the Ningxia region, which is associated with the consumption of undercooked contaminated pork [29, 33]. It is potential for *Y. enterocolitica* to contaminate carcasses and pork products during the slaughter process [34]. Cross-contamination of other foods is possible when handling and preparing raw pork [35]. Because the transmission of *Y. enterocolitica* involves multiple pathways, the health department should continuously monitor and strengthen epidemiological investigation and training; the animal health department should improve quarantine and hygiene supervision of pigs and regulate the operation of workers in the breeding industry; the environment department should improve the supervision of breeding areas and their surrounding environment; the food regulatory department should continuously focus on and strengthen the testing of food with pig origin; the public should be concerned about food safety and avoid eating raw or undercooked meat products.

Public health and epidemiology face increasing challenges in understanding how environmental trends affect zoonotic disease dynamics [36]. An effective approach to zoonotic disease prevention and control that integrates animal, human, and environmental ecological principles is required [37]. The distribution of *Y. enterocolitica* was influenced by a variety of factors. Correlation analyses revealed that altitude, mean temperature and precipitation were important factors influencing the epidemiological diversity of *Y. enterocolitica* in animals in China, and were strongly related to climatic and agricultural



characteristics [14]. It was evident that temperature and precipitation of the collection locations were closely related to the biotype, serotype, and pathogenicity of the isolates. The mean temperature and mean precipitation for pathogenic isolates were 12.70 °C and 1.34 mm, respectively, whereas the mean temperature and mean precipitation of non-pathogenic isolates were 16.85 °C and 2.54 mm, respectively, with the differences being highly statistically significant. This was because that *Y. enterocolitica* was a psychrotrophic bacteria that grew more slowly than most Gram-negative bacteria at room temperature and was unable to compete in mixed populations at high temperatures, but was able to grow, multiply and survive for long periods at low temperatures [38]. *Y. enterocolitica* must acclimate to the temperature of the human host before infection, therefore temperature-dependent chromosomal and plasmid-associated virulence determinants play an essential function [39]. In terms of precipitation, Ningxia has a typical continental semi-humid and semi-arid climate, with the rainy season lasting from June to September. Therefore we hypothesized that the significant differences in pathogenicity and precipitation were directly related to temperature. We used both unifactor and multifactor analysis in the prediction model, and it was obvious that the multifactor prediction result (temperature + precipitation + host) was the most optimal. Because neither humans nor microorganisms exist as independent individuals, various environment factors inevitably influence how organisms survive and contribute to the formation of specific ecological niches, necessitating a multifactorial approach. The random forest model outperformed the other three models in both the training and validation sets. This could benefit from the algorithmic advantages of the random forest, such as high performance and generalisation ability. Accuracy can be maintained even in the presence of missing data. Predictive models are powerful tools for public health and epidemiology, predicting the spread and spillover of pathogenic bacteria based on external environmental factors like temperature, precipitation, altitude, and vegetation distribution, as well as the characteristics of isolates like serotypes and biotypes, and monitoring and alerting on relevant zoonotic diseases.

Through robust and reproducible comparisons, WGS provides unprecedented resolution for characterising *Y. enterocolitica* [40], allowing it to be applied to public health management as well as real-time surveillance and outbreak investigation of foodborne diseases [41]. The ML tree of *Y. enterocolitica* isolates demonstrated that clustering was tightly linked to the biotype, which was consistent with previous findings [42, 43]. Biotype 3, 4, and 5 isolates displayed tight clusters with short terminal branches compared with biotype 1A isolates. This

clade represented the earliest diverging branches of *Y. enterocolitica*, supporting the hypothesis that pathogenic members evolved from non-pathogenic ancestors [42]. Genome analysis reveals different evolutionary processes in pathogens that potentially influence their lifestyles and contribute to specific ecological niches. The typing results revealed regional differences in the distribution of STs and CTs, biotype 1A isolates contained 46 STs (85.2%, 46/54) and 73 (58.4%, 73/125) CTs, indicating a high degree of heterogeneity. Notably, isolates from patients in this study were classified into the same microclades as isolates from food and pigs, implying that *Y. enterocolitica* had the potential to infect humans via pigs and food. According to Karlsson et al. [44], an outbreak of *Y. enterocolitica* 4/O:3 in fresh prewashed spinach occurred in Sweden in 2019. This was completely consistent with the dominant bioserotype of the isolates in this study. Some studies have shown that this biological serotype is associated with approximately 80% of human infections of *Y. enterocolitica* worldwide [45]. Despite the fact that no *Yersinia* outbreaks have been reported in Ningxia, it is worthwhile to raise the alarm.

The distribution of virulence genes was closely related to biotype in the current study. The presence of the virulence factor for iron uptake is attributed to the high pathogenicity of biotype 1B isolates. Iron supply and production of the siderophore transport system were significant factors in the infection of *Y. enterocolitica* [46]. Except for biotype 1B isolates, other pathogenic strains utilise exogenous iron carriers such as ferroflavin B and ferroflavin E [47]. Different pathways of iron uptake may explain differences in the virulence among biotypes [48]. In comparison to biotype 1A isolates, biotypes 2, 3, 4, and 5 isolates had virulence genes associated with the invasion and T3SS. T3SS is primarily composed of multiple *Yersinia* outer proteins (Yops), low calcium response-stimulating proteins (LCRS), and heat-stable *enterocolitica* toxin (Yst) [49]. Gram-negative pathogens inject virulence proteins into the cytoplasm of target eukaryotic cells via Type III secretion system (T3SS) [50]. Biotype 1A strains are regarded as non-pathogenic. However, increasing epidemiological and experimental evidence indicates that biotype 1A strains can cause gastrointestinal disorders [51] and reactive arthritis [52] as well as being implicated in foodborne and nosocomial outbreaks around the world [53], implying a role for the *ystB* gene in the pathogenic process [54]. In the present research, 88.8% (n=79) of the Biotype 1A isolates possessed the *ystB* gene. Although there was no evidence that biotype 1A isolates were pathogenic in this study, researchers should remain vigilant for non-pathogenic *Y. enterocolitica* as well.



There were several limitations to our study. Firstly, the patient sample size was small, and it was collected in 2018 and 2019. The reason is that symptom of diarrhoea caused by *Y. enterocolitica* are usually mild and self-limiting. Patients rarely visit outpatient clinics. Secondly, the collection of samples from both food and patient sources lacked quantitative homogeneity and temporal consistency. Therefore, statistical data and predictive models may be affected. Thirdly, we used ecological epidemiological study to explore the relationship between macro-level environmental factors and pathogenicity of *Y. enterocolitica*. Ecological fallacies in ecological studies are the most significant drawback of ecological studies and are difficult to avoid. Finally, there is a scarcity of epidemiologically relevant information on isolates from diarrhoea patients. These constraints emphasized the significance of the One Health approach. Future work in the human, animal, environmental, and other relevant sectors should be followed by the Tripartite Zoonoses Guide to optimally support countries and regions in addressing zoonotic diseases through a multisectoral, One Health approach [55]. Concrete measures such as early warning and early detection through the Global Early Warning System for health threats and emerging risks at the human-animal-ecosystems interface (GLEWS+), linking surveillance systems, establishing standardised detection methods, sharing laboratory resources, improving laboratory capacity, expanding epidemiology training, and conducting joint sectoral risk assessments [56, 57]. One Health offers a promising and valuable approach to successfully achieving optimal human and animal health while considering the protection of ecosystems and our shared environment.

## Conclusions

The distribution of STs and CTs was related to biotype, geographical area, serotype, and host. The distribution of virulence factors was closely related to the biotype. The distribution of pathogenic isolates was correlated to the environment factors of temperature and precipitation, and machine learning models predicted that the epic of *Y. enterocolitica* areas in the Ningxia region was concentrated in the north, northwest, and south. Our research integrated the ecological modeling and molecular epidemiology study which lead to the public health of *Y. enterocolitica*. It also served as a guide for epidemiological and public health analyses. Zoonotic diseases and other public safety threats can be better addressed using this approach, which addresses health threats at the human-animal-environment interface in a synergistic manner.

## Abbreviations

ANI	Average nucleotide identity
SNP	Single nucleotide polymorphism
cgMLST	Core gene multilocus sequence typing
RF	Random forest
STs	Sequence types
CTs	CgMLST types
NDVI	Normalized difference vegetation index
ML	Maximum likelihood
VFDB	Virulence factor database
HeirCC	Hierarchical cgMLST clustering
MST	Minimum spanning tree
NJ	Neighbor-joining
GLM	General linear model
RF	Random forest model
XGBOOST	eXtreme gradient boosting
ROC	Receiver operating characteristic
AUC	Area under the curves
BAPS	Bayesian analysis of population structure
HC	Hierarchical clustering
Yops	<i>Yersinia</i> outer proteins
LCRS	Low calcium response-stimulating proteins
Yst	Heat-stable <i>enterocolitica</i> toxin
T3SS	Type III secretion system
GLEWS+	The Global Early Warning System for health threats and emerging risks at the human-animal-ecosystems interface

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40249-023-01063-6>.

**Additional file 1: Fig. S1** Differences in biotype of *Y. enterocolitica* isolates and ecological factors. a Temperature. b Precipitation. c Altitude. d NDVI.

**Additional file 2: Fig. S2** Differences in serotype of *Y. enterocolitica* isolates and ecological factors. a Temperature. b Precipitation. c Altitude. d NDVI.

**Additional file 3: Fig. S3** Differences in host of *Y. enterocolitica* isolates and ecological factors. a Temperature. b Precipitation. c Altitude. d NDVI.

**Additional file 4: Fig. S4** Variable importance of predictive model for pathogenicity of *Y. enterocolitica*. a GLM (General Linear Model). b RF (Random Forest model). c XGBOOST (eXtreme Gradient Boosting).

**Additional file 5: Fig. S5** The risk of pathogenic of *Yersinia enterocolitica* by machine learning model predictions in study regions. a–c The RF model predict the probability of pathogenic of *Y. enterocolitica* with sample source from animal, food, and human, respectively. d–f The GLM model predict the probability of pathogenic of *Y. enterocolitica* with sample source from animal, food, and human, respectively. g–i The XGBOOST model predict the probability of pathogenic of *Y. enterocolitica* with sample source from animal, food, and human, respectively. The ecological variables of temperature and precipitation were based on the monthly mean value of Ningxia region on August 2019.

**Additional file 6: Fig. S6** The Minimum Spanning Trees (MST) of *Y. enterocolitica* isolates. The circle size was proportional to the number of isolates. Links between circles were represented according to the number of allelic differences between STs. a Biotype. b Serotype.

**Additional file 7: Fig. S7** The Neighbor-joining (NJ) tree of *Y. enterocolitica* isolates based on cgMLST. a Biotype. b Serotype. c Host. d Country.

**Additional file 8: Table S1.** Summary of environment and climatic data sources. **Table S2.** List of the *Yersinia* reference genomes. **Table S3.** List of 88 *Y. enterocolitica* strains downloaded from the NCBI database. **Table S4.** Analysis results of training and test datasets. **Table S5.** List of 270 *Yersinia* isolates in the Ningxia Hui Autonomous Region **Table S6.** Origin, serotype, STs, and CTs of 187 strains of *Y. enterocolitica*.

### Acknowledgements

We thank all the staff of the Bacteriology Section of the Ningxia Hui Autonomous Region Centre for Disease Control and Prevention for their help in collecting the isolates.

### Author contributions

YY performed the preparation of DNA samples, identification of biotypes, clustering analysis, cgMLST analysis; YY and ZJX drafted the manuscript; ZJX, ZSX, and LY performed the ecological and machine learning analysis; SM, LX, and HQ mainly performed the strain isolation; XS performed the data analysis; LZG, LY, JHQ, and HXX revised the manuscript; ZXZ and LZJ participated in the design of this study and reviewed the manuscript; LZJ managed the project. All authors read and approved the final manuscript.

### Funding

This work was supported by grants from the National Key R&D Program of China (2019YFC1200700, 2019YFC1200601-6), National Natural Fund Projects (32060816), and the Ningxia Hui Autonomous Region Natural Genetic Project (2022AAC03719).

### Availability of data and materials

FASTQ datafiles have been deposited in NCBI-SRA and public archives under the project's accession number PRJNA993921 (Centers for Disease Control and Prevention). The accession numbers of all isolates are provided in Additional file 8: Table S5. All data generated or analyzed during this study are included in this published article. The Additional information files are freely available to any scientist wishing to use them for non-commercial purposes upon request via e-mail with the corresponding author.

### Declarations

#### Ethics approval and consent to participate

This study was carried out according to the principles of the Declaration of Helsinki, the ARRIVE guidelines and the 3Rs. The research protocol was approved by the Ethics Committees of the Chinese Center for Disease Control and Prevention (ICDC-2019015). All isolates from humans were collected as part of a standard clinical investigation of patients with diarrhea. These patients were anonymous and had signed an informed consent form. All animal samples used in this study were approved by the Research Animal Ethics and Ethical Committee of the Chinese Centre for Disease Control and Prevention (2020-025). Written informed consent was obtained from the owners for the participation of their animals in this study. The *Yersinia* isolates used in this study were obtained from academic articles published in PubMed.

#### Consent for publication

Not applicable.

#### Competing interests

No potential conflict of interest was reported by the author(s).

#### Author details

<sup>1</sup>Key Laboratory of the Ministry of Education for the Conservation and Utilization of Special Biological Resources of Western China, Ningxia University, Yinchuan, People's Republic of China. <sup>2</sup>State Key Laboratory for Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, People's Republic of China. <sup>3</sup>Ningxia Hui Autonomous Region Food Testing and Research Institute, Yinchuan, People's Republic of China. <sup>4</sup>Department of Nephrology, Ruijin Hospital, Institute of Nephrology, Shanghai Jiao Tong University School of Medicine, Shanghai, People's Republic of China. <sup>5</sup>School of Global Health, Chinese Center for Tropical Diseases Research-Shanghai Jiao Tong University School of Medicine, Shanghai, People's Republic of China. <sup>6</sup>Ningxia Hui Autonomous Region Centre for Disease Control and Prevention, Yinchuan, People's Republic of China. <sup>7</sup>Longhua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, People's Republic of China. <sup>8</sup>Department of Computer Science, Hong Kong Baptist University, Hong Kong, Special Administrative Region, People's Republic of China.

Received: 7 August 2022 Accepted: 5 February 2023

Published online: 21 April 2023

### References

- Ahmed HA, Tahoun A, Elez RA, El-Hamid MA, Ellatif SA. Prevalence of *Yersinia enterocolitica* in milk and dairy products and the effects of storage temperatures on survival and virulence gene expression. *Int Dairy J*. 2019;94:16–21.
- Bari ML, Hossain MA, Isshiki K, Ukuku D. Review article behavior of *Yersinia enterocolitica* in foods. *J Pathog*. 2011;420732:1–13.
- Valentin-Weigand P, Heeseemann J, Dersch P. Unique virulence properties of *Yersinia enterocolitica* O:3—an emerging zoonotic pathogen using pigs as preferred reservoir host. *Int J Med Microbiol*. 2014;304(7):824–34.
- European Food Safety Authority corresponding author and European Centre for Disease Prevention and Control. The European Union One Health 2019 Zoonoses Report. EFSA J. 2021;19(2): e06406.
- The Institute of Environmental Science and Research Ltd. <https://www.esr.cri.nz/home/about-esr/our-science-in-action/detection-of-patho-yersinia-in-foods/>. Accessed 6 Nov 2018.
- Wang X, Qiu H, Dong J, Cui ZG, Kan B, Xiao YC, et al. O:8 serotype *Yersinia enterocolitica* strains in China. *Int J Food Microbiol*. 2008;125(3):259–66.
- Guo BC, Liu X, Hao Q, Yan L, Xie MY, Xue XY, et al. Monitoring and analysis of *Yersinia enterocolitica* in Ningxia area. *Mod Med Health*. 2011;27(018):2724–6.
- Peng ZX, Zou MY, Xu J, Guan WY, Li Y, Liu DR, et al. Antimicrobial susceptibility and drug-resistance genes of *Yersinia* spp. of retailed poultry in 4 provinces of China. *Chin J Prev Med*. 2018;52(4):358–63.
- The Steering Committee by the Standards Unit. UK Standards for Microbiology Investigations B 30: Investigation of faecal specimens for enteric pathogens. Public Health England. 2019.
- Strydom H, Wang J, Paine S, Dyet K, Cullen K, Wright J. Evaluating sub-typing methods for pathogenic *Yersinia enterocolitica* to support outbreak investigations in New Zealand. *Epidemiol Infect*. 2019;147: e186.
- Jones BA, Grace D, Kock R, Alonso S, Rushton J, Said MY, et al. Zoonosis emergence linked to agricultural intensification and environmental change. *Proc Natl Acad Sci U S A*. 2013;110(21):8399–404.
- Koelle K, Rodó X, Pascual M, Yunus M, Mostafa G. Refractory periods and climate forcing in cholera dynamics. *Nature*. 2005;436(7051):696–700.
- Tian H, Yu P, Bjørnstad ON, Cazelles B, Yang J, Tan H, et al. Anthropogenically driven environmental changes shift the ecological dynamics of hemorrhagic fever with renal syndrome. *PLoS Pathog*. 2017;13(1): e1006198. <https://doi.org/10.1371/journal.ppat.1006198>.
- Liang J, Duan R, Xia S, Hao Q, Yang JC, Xiao YC, et al. Ecology and geographic distribution of *Yersinia enterocolitica* among livestock and wildlife in China. *Vet Microbiol*. 2015;178(1–2):125–31.
- Duan R, Liang J, Shi G, Cui ZG, Hai R, Wang P, et al. Homology analysis of pathogenic *Yersinia* Species *Yersinia enterocolitica*, *Yersinia pseudotuberculosis*, and *Yersinia pestis* based on multilocus sequence typing. *J Clin Microbiol*. 2014;52(1):20–9.
- Stamm I, Hailer M, Depner B, Kopp PA, Rau J. *Yersinia enterocolitica* in diagnostic fecal samples from European dogs and cats: identification by fourier transform infrared spectroscopy and matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J Clin Microbiol*. 2013;51(3):887–93.
- Wang X, Cui Z, Jin D, Tang L, Xia S, Wang H, et al. Distribution of pathogenic *Yersinia enterocolitica* in China. *Eur J Clin Microbiol Infect Dis*. 2009;28(10):1237–44.
- Bottone EJ. *Yersinia enterocolitica*: the charisma continues. *Clin Microbiol Rev*. 1997;10(2):257–76.
- Xiao YC, Wang X, Qiu HY, et al. Study of Biotyping for pathogenic *Yersinia enterocolitica* strains in China. *Chin J Zoonoses*. 2010;26(07):651–53.
- Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, et al. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics*. 2009;25(15):1966–7.
- Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun*. 2018;9(1):5114.

22. Delcher AL, Salzberg SL, Phillippy AM. Using MUMmer to Identify Similar Regions in Large Sequence Sets. *Curr Protoc Bioinform.* 2003;Chapter 10:Unit 10.3. <https://doi.org/10.1002/0471250953>.
23. Gary B. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 1999;27(2):573–80.
24. Ivica L, Peer B. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 2016;44(W1):W242–5.
25. Zhou Z, Alikhan NF, Mohamed K, Fan Y, Achtman M. The Enterobase user's guide, with case studies on *Salmonella* transmissions, *Yersinia pestis* phylogeny and *Escherichia* core genomic diversity. *Genome Res.* 2020;30(1):138–52.
26. Hall M, Chattaway MA, Reuter S, Savin C, McNally A. Use of whole-genus genome sequence data to develop a multilocus sequence typing tool that accurately identifies *Yersinia* isolates to the species and subspecies levels. *J Clin Microbiol.* 2015;53(1):35–42.
27. Achtman M. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. *Proc Natl Acad Sci U S A.* 1999;96(24):14043–8.
28. Rivas L, Strydom H, Paine S, Wang J, Wright J. Yersiniosis in New Zealand. *Pathogens.* 2021;10(2):191.
29. Martins BTF, Botelho CV, Silva DAL, Lanna FGPA, Grossi JL, Campos-Galvão MEM, et al. *Yersinia enterocolitica* in a Brazilian pork production chain: tracking of contamination routes, virulence and antimicrobial resistance. *Int J Food Microbiol.* 2018;276:5–9.
30. Bancercz-Kisiel A, Szweda W. Yersiniosis—zoonotic foodborne disease of relevance to public health. *Ann Agric Environ Med.* 2015;22(3):397–402.
31. Fosse J, Seegers H, Magras C. Prevalence and risk factors for bacterial food-borne zoonotic hazards in slaughter pigs: a review. *Zoonoses Public Health.* 2009;56(8):429–54.
32. Drummond N, Murphy BP, Ringwood T, Prentice MB, Buckley JF, Faning S. *Yersinia enterocolitica*: a brief review of the issues relating to the zoonotic pathogen, public health challenges, and the pork production chain. *Foodborne Pathog Dis.* 2012;9(3):179–89.
33. Cressey PJ, Lake RJ, Thornley C, Campbell D. Expert elicitation for estimation of the proportion foodborne for selected microbial pathogens in New Zealand. *Foodborne Pathog Dis.* 2019;16(8):543–9.
34. Van Damme I, Berkvens D, Vanantwerpen G, Bare J, Houf K, Wauters G, et al. Contamination of freshly slaughtered pig carcasses with enteropathogenic *Yersinia* spp.: distribution, quantification and identification of risk factors. *Int J Food Microbiol.* 2015;204:33–40.
35. Boqvist S, Pettersson H, Svensson A, Andersson Y. Sources of sporadic *Yersinia enterocolitica* infection in children in Sweden, 2004: a case-control study. *Epidemiol Infect.* 2009;137(6):897–905.
36. Altizer S, Ostfeld RS, Johnson PT, Kutz S, Harvell CD. Climate change and infectious diseases: from evidence to a predictive framework. *Science.* 2013;341(6145):514–9.
37. Johnson PT, de Roode JC, Fenton A. Why infectious disease research needs community ecology. *Science.* 2015;349(6252):1259504.
38. Schiemann DA. *Yersinia enterocolitica* in milk and dairy products. *J Dairy Sci.* 1987;70:383–91.
39. Bottone EJ. *Yersinia enterocolitica*: overview and epidemiologic correlates. *Microbes Infect.* 1999;1(4):323–33.
40. Inns T, Flanagan S, Greig DR, Jenkins C, Seddon K, Chin T, Cartwright J. First use of whole-genome sequencing to investigate a cluster of *Yersinia enterocolitica*, Liverpool, United Kingdom, 2017. *J Med Microbiol.* 2018;67(12):1747–52.
41. Technical Background Paper: Applications of Whole Genome Sequencing in Food Safety Management; Food and Agriculture Organization of the United Nations: Rome, Italy. 2016. Food and Agriculture Organization of the United Nations. <http://www.fao.org/3/a-i5619e.pdf>. Accessed 14 Jan 2021.
42. Reuter S, Connor TR, Barquist L, Walker D, Thomson NR. Parallel independent evolution of pathogenicity within the genus *Yersinia*. *Proc Natl Acad Sci U S A.* 2014;111(18):6768–73.
43. Savin C, Criscuolo A, Guglielmini J, Guerni ASL, Carniel E, Pizarro-Cerdá J, et al. Genus-wide *Yersinia* core-genome multilocus sequence typing for species identification and strain characterization. *Microb Genom.* 2019;5(10): e000301.
44. Karlsson PA, Tano E, Jernberg C, Hickman RA, Guy L, Järhult JD, et al. Molecular characterization of multidrug-resistant *Yersinia enterocolitica* from foodborne outbreaks in Sweden. *Front Microbiol.* 2021;12: 664665.
45. Duan R, Liang J, Zhang J, Chen Y, Wang J, Tong J, et al. Prevalence of *Yersinia enterocolitica* Bioserotype 3/O:3 among children with diarrhea, China, 2010–2015. *Emerg Infect Dis.* 2017;23(9):1502–9.
46. Huang Y, Wang X, Cui Z, Yang Y, Xiao Y, Tang L, et al. Possible use of ail and foxa polymorphisms for detecting pathogenic *Yersinia enterocolitica*. *BMC Microbiol.* 2010;10:211.
47. Baumler A, Koebnik R, Stojiljkovic I, Heesemann J, Braun V, Hantke K. Survey on newly characterized iron uptake systems of *Yersinia enterocolitica*. *Zentralbl Bakteriol.* 1993;278(2–3):416–24.
48. Bakour R, Balligand G, Laroche Y, Cornelis G, Wauters G. A simple adult-mouse test for tissue invasiveness in *Yersinia enterocolitica* strains of low experimental virulence. *J Med Microbiol.* 1985;19(2):237–46.
49. Ramamurthy KS, Schneewind O. Type III protein secretion in *Yersinia* Specie. *Annu Rev Cell Dev Biol.* 2002;18:107–33.
50. Fàbrega A, Vila J. *Yersinia enterocolitica*: pathogenesis, virulence and antimicrobial resistance. *Enferm Infecc Microbiol Clin.* 2012;30(1):24–32.
51. Tennant SM, Grant TH, Robins-Browne RM. Pathogenicity of *Yersinia enterocolitica* biotype 1A. *FEMS Immunol Med Microbiol.* 2003;38(2):127–37.
52. Tuompo R, Hannu T, Huovinen E, Sihvonen L, Siitonen A, Leirisalo-Repo M. *Yersinia enterocolitica* biotype 1A: a possible new trigger of reactive arthritis. *Rheumatol Int.* 2017;37(11):1863–9.
53. Bhagat N, Viridi JS. The enigma of *Yersinia enterocolitica* biovar 1A. *Crit Rev Microbiol.* 2011;37(1):25–39.
54. Ramamurthy T, Yoshino K, Huang X, Balakrish Nair G, Carniel E, Maruyama T, et al. The novel heat-stable enterotoxin subtype gene (*ystB*) of *Yersinia enterocolitica*: nucleotide sequence and distribution of the *yst* genes. *Microb Pathog.* 1997;23(4):189–200.
55. Taking a multisectoral, one health approach: a tripartite guide to addressing zoonotic diseases in countries. 2019. World Health Organization, Food and Agriculture Organization of the United Nations & World Organisation for Animal Health. <https://apps.who.int/iris/handle/10665/325620>. Accessed 28 Jun 2019.
56. Behravesh CB. Introduction. One Health: over a decade of progress on the road to sustainability. *Rev Sci Tech.* 2019;38(1):21–50.
57. Sinclair JR. Importance of a One Health approach in advancing global health security and the sustainable development goals. *Rev Sci Tech.* 2019;38(1):145–54.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

