# Supplementary tables and figures

Table S1 Number of distinct C. coli multi-locus sequence types (STs) and isolates from humans and food sources. Percentages, where given, denote of the fraction of all C. coli STs/isolates with the same source (humans, chicken, pig, or ruminant).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Origin | STs | Isolates | Unique\* STs (%) | Isolates from Unique\* STs (%) | STs without cases (%) | Isolates from STs without cases (%) |
| Human | 31 | 125 | 18 (58%) | 19 (15%) | N/A | N/A |
| Chicken | 40 | 283 | 29 (73%) | 101 (36%) | 28 (70%) | 73 (26%) |
| Pig | 23 | 28 | 15 (65%) | 18 (64%) | 19 (83%) | 23 (82%) |
| Ruminant | 7 | 16 | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) |

\*Unique types for sources are types that were found in that source and no other source (but potentially in humans), while unique types for humans are those found in cases but not in any of the three sources.

Table S2 Number of distinct C. jejuni multi-locus sequence types (STs) and isolates from humans and food sources. Percentages, where given, denote of the fraction of all C. jejuni STs/isolates with the same source (humans, chicken, pig, or ruminant).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Origin | STs | Isolates | Unique\* STs (%) | Isolates from Unique\* STs (%) | STs without cases (%) | Isolates from STs without cases (%) |
| Human | 87 | 585 | 48 (55%) | 115 (20%) | N/A | N/A |
| Chicken | 48 | 197 | 31 (65%) | 91 (46%) | 10 (21%) | 17 (8.6%) |
| Pig | 10 | 16 | 2 (20%) | 2 (13%) | 2 (20%) | 2 (13%) |
| Ruminant | 18 | 72 | 3 (17%) | 3 (4.2%) | 2 (11%) | 2 (2.8%) |

\*Unique types for sources are types that were found in that source and no other source (but potentially in humans), while unique types for humans are those found in cases but not in any of the three sources.

Table S3 Comparison of observed and estimated relative abundance (%) of two multi-locus sequence types (STs) in the three sources and ‘unsampled source’ (where relevant). ST50 and ST48 were the 1st and 4th most common STs in humans in our study. The relative abundance is the percentage of all Campylobacter isolates from a given source that belong to the indicated sequence type. Observed values are the percentage of isolates of the given sequence type with numbers given in parentheses. The point estimates for the models are the posterior mean with 95% credible intervals.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model1 | Sequence Type 50 | | | | Sequence Type 48 | | | |
| Ruminant | Pig | Chicken | Unsampled | Ruminant | Pig | Chicken | Unsampled |
| Observed | 24  (21/88) | 11 (5/44) | 5.2 (25/480) | NA  (0/0) | 1.1 (1/88) | 0.0 (0/44) | 0.8 (4/480) | NA  (0/0) |
| M1 | 8.4  (5.4-12.0) | 2.8 (1.1-5.4) | 4.1 (2.8-5.8) | 0.6 (1.8x10-2-2.2) | 0.8 (0.1-2.2) | 0.5 (1.1x10-02-1.7) | 1.0 (0.4-1.8) | 0.7 (1.8x10-2-2.5) |
| M2 | 8.4  (5.3-12.2) | 2.8 (1.0-5.1) | 4.1 (2.8-5.8) | NA | 0.8 (0.1-2.2) | 0.5 (1.2x10-2-1.8) | 1.0 (0.4-1.8) | NA |
| M3 | 20  (13-28) | 8.6 (2.9-17) | 5.2 (3.3-7.3) | 0.8 (2.3x10-16-8.3) | 1.4 (0.1-4.5) | 0.2 (6.9x10-17-1.8) | 1.0 (0.3-2.0) | 2.8 (7.1x10-14-13.9) |
| M4 | 19.3  (13-27) | 8.5 (2.8-17) | 5.1 (3.4-7.2) | NA | 1.5 (0.1-4.5) | 0.2 (2.0x10-16-2.0) | 0.9 (0.3-2.0) | NA |
| M5 & M8 | 8.4  (5.3-12) | 2.7 (1.0-5.3) | 4.0 (2.6-5.6) | 0.6 (1.5x10-2-2.1) | 0.8 (0.1-2.1) | 0.5 (1.2x10-2-1.7) | 0.8 (0.2-1.6) | 0.6 (1.5x10-2-2.1) |
| M6 & M9 | 8.4  (5.3-12) | 2.7 (1.0-5.3) | 4.0 (2.6-5.6) | NA | 0.8 (0.1-2.1) | 0.5 (1.2x10-2-1.7) | 0.8 (0.2-1.6) | NA |
| M7 & M10 | 16.6 (15-18) | 8.8 (7.3-11) | 4.7 (4.4-5.1) | NA | 0.8 (0.8-0.9) | 0.3 (0.2-0.4) | 0.7 (0.6-0.7) | NA |

1 The pairs of models M5 & M8, M6 & M9, and M7 & M10 use the same model for relative abundance of sequence types in sources. See Table 1 for a full description of the models.

Table S4 Comparison of observed and estimated relative abundance (%) of multi-locus sequence types (STs) in the three sources and ‘unsampled source’ (where relevant). The relative abundance is the percentage of all Campylobacter isolates from a given source that belong to the indicated sequence type. Observed values are the percentage of isolates of the given sequence type with numbers given in parentheses. The point estimates for the models are the posterior mean with 95% credible intervals.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model1 | Sequence Type 827 | | | | Sequence Type 2398 | | | |
| Ruminant | Pig | Chicken | Unsampled | Ruminant | Pig | Chicken | Unsampled |
| Observed | 9.1  (8/88) | 4.5 (2/44) | 9.6 (46/480) | NA  (0/0) | 0.0 (0/88) | 0.0 (0/44) | 0.0 (0/480) | NA  (0/0) |
| M1 | 3.4 (1.5-6.0) | 1.4 (0.3-3.2) | 7.1 (5.2-9.2) | 0.6 (1.6x10-2-2.0) | 0.4 (1.2x10-2-1.6) | 0.5 (1.3x10-2-1.9) | 0.3 (1.4x10-2-0.9) | 0.9 (3.0x10-2-2.9) |
| M2 | 3.4 (1.6-5.8) | 1.4 (0.3-3.3) | 7.1 (5.2-9.3) | NA | 0.5 (1.5x10-2-1.7) | 0.6 (1.7x10-2-2.0) | 0.3 (2.7x10-2-0.9) | NA |
| M3 | 7.1 (3.2-12.3) | 3.5 (0.5-9.4) | 9.2 (6.8-12.0) | 0.5 (1.0x10-16-4.3) | 0.2 (1.3x10-16-1.9) | 0.2 (1.4x10-16-2.4) | 2.9x10-2 (1.3x10-16-0.3) | 5.0 (3.9x10-6-15.3) |
| M4 | 7.0 (3.2-12.2) | 3.4 (0.4-9.2) | 9.2 (6.9-11.9) | NA | 0.9 (1.3x10-13-3.2) | 0.7 (1.1x10-15-4.9) | 0.1 (4.6x10-17-0.5) | NA |
| M5 & M8 | 3.4 (1.6-5.9) | 1.4 (0.3-3.3) | 7.2 (5.3-9.3) | 0.6 (1.5x10-2-2.1) | 0.4 (9.7x10-3-1.4) | 0.5 (1.2x10-2-1.7) | 0.2 (3.9x10-3-0.6) | 0.6 (1.5x10-2-2.1) |
| M6 & M9 | 3.4 (1.6-5.9) | 1.4 (0.3-3.3) | 7.2  (5.3-9.3) | NA | 0.4 (9.7x10-3-1.4) | 0.5 (1.2x10-2-1.7) | 0.2 (3.9x10-3-0.6) | NA |
| M7 & M10 | 7.3 (7.0-7.6) | 5.7  (5.0-6.3) | 7.8 (7.5-8.0) | NA | 2.1x10-11 (8.9x10-12-4.4x10-11) | 5.7x10-12 (1.4x10-12-1.3x10-11) | 1.5x10-11 (6.8x10-12-3.1x10-11) | NA |

1 The pairs of models M5 & M8, M6 & M9, and M7 & M10 use the same model for relative abundance of sequence types in sources. See Table 1 for a full description of the models.

Table S5 Predictive performance of ten source attribution models, measured with Watanabe-Akaike Information Criterion (WAIC) for the prediction of the relative abundance of sequence types in cases (all models: M1-M10) or cases and sources (jointly estimated models only: M1-M4). Standard errors are given in parentheses. WAIC is smaller in models with better fit. WAIC difference is calculated with respect to best model (M1 for predictions in cases, M3 for predictions in cases and sources). However, we only consider differences substantive when greater than five times the standard error for the difference, e.g. when considering predictions in cases, M1 is substantively better than all models except M2 and M7.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Cases | | | Cases and Sources | | |
| Model | WAIC | WAIC difference | Effective # parameters | WAIC | WAIC difference | Effective # parameters |
| M1 | 5,819.2 (76.4) | 0.0 (0.0) | 102.8 (5.6) | 10,785.4 (95.8) | 112.7 (24.6) | 209.6 (7.5) |
| M2 | 5,822.2 (76.8) | 3.0 (1.1) | 103.5 (5.7) | 10,792.0 (96.3) | 119.4 (24.1) | 210.1 (7.6) |
| M3 | 5,870.4 (82.5) | 51.1 (9.6) | 131.2 (9.0) | 10,672.7 (112.1) | 0.0 (0.0) | 312.7 (14.5) |
| M4 | 5,881.2 (84.6) | 62.0 (12.1) | 135.5 (9.4) | 10,708.8 (114.1) | 36.2 (5.1) | 318.7 (14.9) |
| M5 | 5,828.4 (77.5) | 9.2 (1.6) | 109.8 (6.2) | — | — | — |
| M6 | 5,831.1 (78.0) | 11.9 (2.2) | 110.9 (6.3) | — | — | — |
| M7 | 5,896.1 (89.5) | 76.8 (92.4) | 172.8 (13.3) | — | — | — |
| M8 | 6,690.1 (60.4) | 870.8 (51.5) | 108.4 (3.3) | — | — | — |
| M9 | 6,721.2 (63.6) | 902.0 (52.6) | 116.3 (4.1) | — | — | — |
| M10 | 9,950.8 (382.8) | 4,131.5 (361.1) | 17.5 (2.0) | — | — | — |

Table S6 Estimates and 95% credible intervals for selected parameters related to type transmission potential in the seven models (M1-M7) that allowed for transmission potential variability between sequence types. See Table 1 for further details of models.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model† | \* |  |  |  |  |
| M1 | 1.2 (1.0-1.5) | 18 (4.6-85) | 15 (4.7-48) | 12 (5.2-29) | 0.6 (0.3-1.5) |
| M2 | 1.2 (1.0-1.5) | 22 (5.6-110) | 15 (5.6-49) | 11 (4.7-25) | 0.6 (0.3-1.2) |
| M3 | 1.0 (0.6-1.3) | 4.8 (0.8-56) | 9.2 (1.2-43) | 6.0 (2.3-16) | 0.5 (0.2-1.0) |
| M4 | 1.2 (0.9-1.6) | 23 (4.1-210) | 16 (4.2-66) | 5.0 (2.1-13) | 0.3 (0.1-0.6) |
| M5 | 1.4 (1.2-1.7) | 40 (7.9-270) | 19 (6.2-70) | 10 (4.5-25) | 0.7 (0.3-1.8) |
| M6 | 1.4 (1.2-1.7) | 34 (6.6-170) | 17 (5.6-58) | 12 (5.4-29) | 0.7 (0.3-1.4) |
| M7 | 12 (10-13) | 7.6x1011 (2.2x1011-2.5x1012) | 28 (14-61) | 12 (6.4-25) | 3.4x10-2 (4.2x10-2-0.4) |

† Models M8-M10 assumed all types had equal transmission potential. See Table 1 for a full description of models M1-M7.

\*The standard deviation (on the log scale) of the log-normal distribution of type transmission potential parameters .

Venn diagram

Description automatically generated

Figure S1 Venn diagram summarising the number of multi-locus sequence types (STs) found in cases and sources, or combinations of cases and sources . For instance, 66 STs were found only in cases, 26 STs were found in chicken and cases but not in pigs or ruminants, and 8 STs (representing 32% [423/1322] of study isolates) were found in cases and all sources.

Chart

Description automatically generated

Figure S2 Source attribution proportions of four selected multi-locus sequence types (STs) to three sampled sources in ten models (M1-M10, left to right). Four models (M1, M3, M5, M8) also include a fourth, ‘unsampled source’. See Table 1 for more details about the models. ST50 (C. jejuni) was the most observed type in cases, ruminants, and pigs and the 5th most common type in chickens. ST48 (C. jejuni) was 4th most common type in humans, but rare or absent in all sources. ST827 (C. coli) was the 2nd most common type in chickens, and found in the other sources, but relatively uncommon in cases. ST2398 (C. jejuni) was more common in cases than ST827, but not detected in any sources. Estimates of relative abundance of these types in the respective sources can be found in Tables S3 and S4. Estimates of the transmission potential of the four types can be found in Table S6.

Graphical user interface, chart

Description automatically generated

Figure S3 Percent difference in source attribution proportions between urban and rural populations (urban cases as references) for eight models (M1,M2,M5-M10). See Table 1 for model details. Vertical bars are 95% credible intervals (CrI). Note that all CrIs include 0% (no difference).

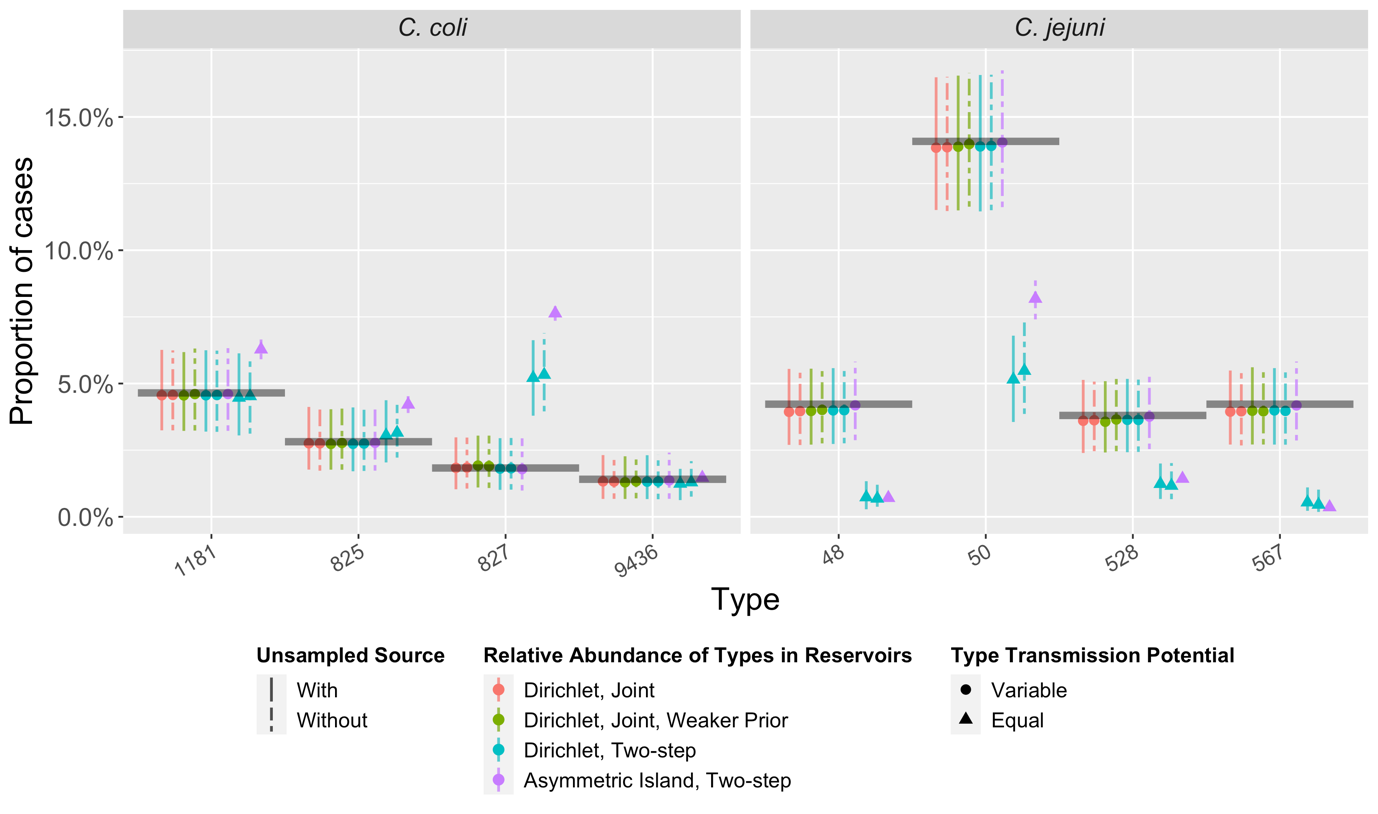


Figure S4 The observed proportion of cases due to the five most common multi-locus sequence types (STs) from C. coli and C. jejuni (black horizontal lines) compared to predictions (coloured points and vertical 95% credible intervals) under 10 source attribution models (M1-M10, left to right). Note that for some STs, the credible intervals for M10 (asymmetric island model) are so narrow that they are not visible.

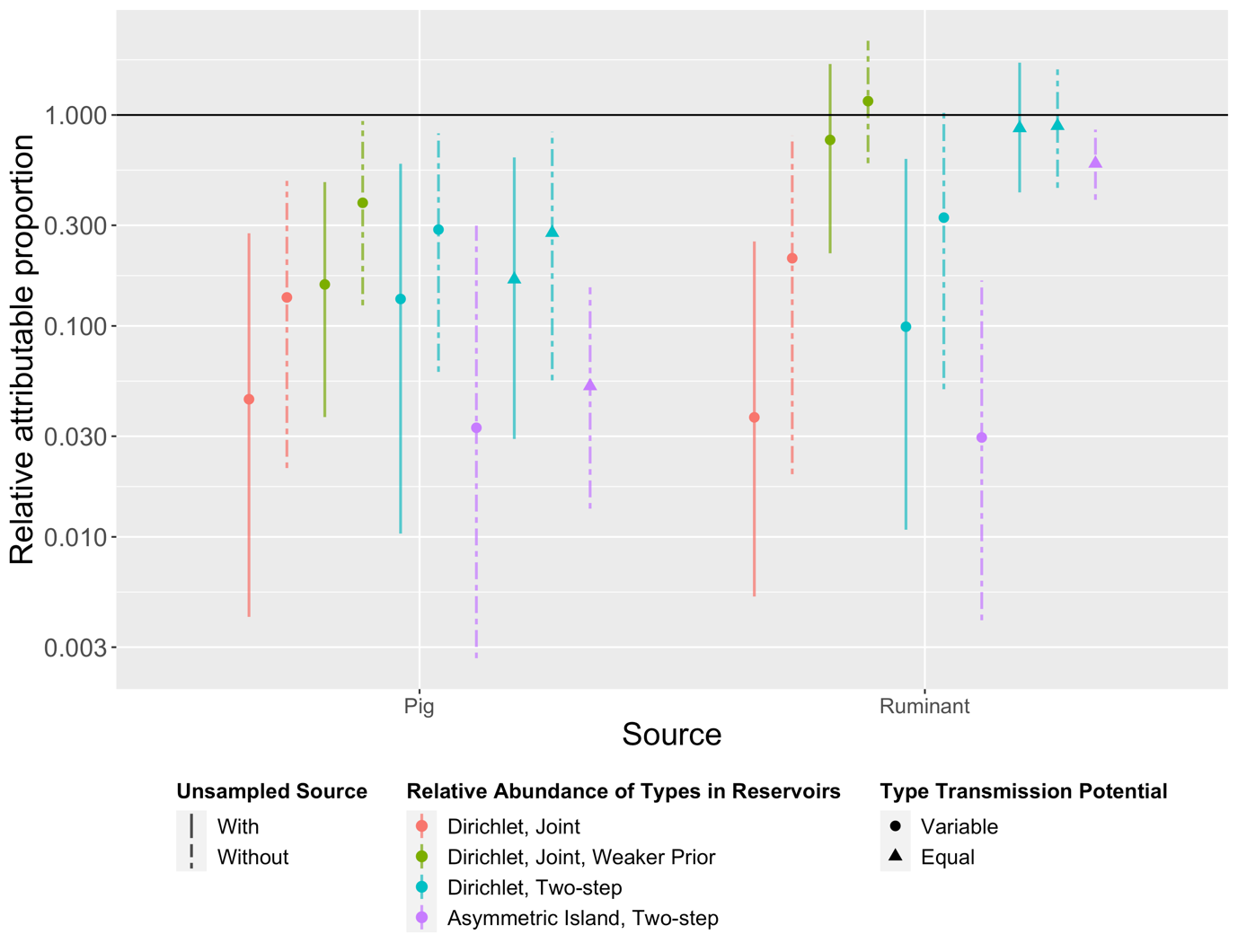


Figure S5 Estimates of relative attributable proportion (RAP) of campylobacteriosis under ten source attribution models (M1-M10, left to right). RAP was calculated by dividing attribution proportion by the annual Australian consumption of meat products derived from that source and normalised against a reference source (chicken). All models indicated pig meat poses less risk (lower RAP) than chicken. Note the y-axis is on a log scale, and that the ‘unsampled source’ is omitted due to lack of respective consumption statistics or appropriate equivalent exposure measure.