Risk factors for infection with *Campylobacter jejuni flaA* genotypes

L. E. UNICOMB[1,2]*, L. C. O'REILLY[3], M. D. KIRK[4], R. J. STAFFORD[5],
H. V. SMITH[6], N. G. BECKER[2], M. S. PATEL[2], G. L. GILBERT[7]
and the Australian Campylobacter Subtyping Study Group†

[1] *OzFoodNet, Hunter New England Population Health, Wallsend, New South Wales, Australia*
[2] *National Centre for Epidemiology and Population Health, Australia National University, Canberra, Australian Capital Territory, Australia*
[3] *Division of Microbiology & Infectious Diseases, PathWest, Nedlands, Western Australia, Australia*
[4] *OzFoodNet, Department of Health and Ageing, Canberra, Australian Capital Territory, Australia*
[5] *OzFoodNet* (*Queensland*), *Communicable Diseases Unit, Queensland Health, Brisbane, Australia*
[6] *Public Health Microbiology, Queensland Health Scientific Services, Coopers Plains, Queensland, Australia*
[7] *Centre for Infectious Diseases & Microbiology, Institute of Clinical Pathology and Medical Research, New South Wales, Australia*

SUMMARY

We aimed to explore *Campylobacter* genotype-specific risk factors in Australia. Isolates collected prospectively from cases recruited into a case-control study were genotyped using *flaA* restriction fragment-length polymorphism typing (*flaA* genotyping). Exposure information for cases and controls was collected by telephone interview. Risk factors were examined for major *flaA* genotypes using logistic and multinomial regression. Five *flaA* genotypes accounted for 325 of 590 (55%) cases – *flaA*-6b ($n=129$), *flaA*-6 ($n=70$), *flaA*-10 ($n=48$), *flaA*-2 ($n=43$), *flaA*-131 ($n=35$). In Australia, infections due to *flaA*-10 and *flaA*-2 were found to be significantly associated with eating non-poultry meat (beef and ham, respectively) in both case-control and inter-genotype comparisons. All major genotypes apart from *flaA*-10 were associated with chicken consumption in the case-control comparisons. Based on several clinical criteria, infections due to *flaA*-2 were more severe than those due to other genotypes. Thus genotype analysis may reveal genotype-specific niches and differences in virulence and transmission routes.

INTRODUCTION

The disease burden of *Campylobacter* infections is considerable in Australia, where it is the most common notifiable disease [1]. *Campylobacter* infections are largely sporadic and estimated to cause about 223 000 cases of gastroenteritis each year [2]. The need for public health intervention is highlighted by the steady rise in notifications in Australia since the early 1990s [1, 3].

Case-control studies of potential risk factors for campylobacteriosis have been undertaken in several countries [4–11] including Australia [12, 13]. Consumption of chicken is the most commonly identified risk factor for sporadic campylobacteriosis. However, this and other statistically significantly risk factors often do not explain the majority of cases

* Author for correspondence : Ms. L. E. Unicomb, National Centre for Epidemiology and Population Health, Australia National University, Canberra, ACT 0200, Australia.
Email : leanne.unicomb@anu.edu.au
† Members of the Australian Campylobacter Subtyping Study Group are given in the Appendix.

[13, 14]. Other sources of apparently sporadic campylobacteriosis are difficult to detect using current case-control study designs.

Further sources of *Campylobacter* infection may be detected if strains with distinct ecologies (including differential survival characteristics, and therefore specific environmental niches or host preference), with varying virulence (causing disease of variable severity) and different transmission routes could be distinguished. Numerous subtyping methods have been developed and applied to *Campylobacter* isolates [15–19]. These methods, used to determine genomic diversity, generally have been applied to diverse collections of isolates [20, 21]. Methods such as serotyping, multilocus sequence typing (MLST) and pulsed-field gel electrophoresis (PFGE) have been used to distinguish outbreak from sporadic isolates [18, 22, 23]. In addition, a variety of methods have been used to compare animal and human isolates. Some studies have concluded that finding the same subtype in animal and human isolates is evidence of transmission from animals to humans or is evidence of a common source for animals and humans [17, 24–28]. However, the potential to identify genotype-specific risk factors for sporadic infection is yet to be evaluated.

The aim of this study was to determine sources and risk factors of *Campylobacter jejuni* for specific *flaA* genotypes. We hypothesize that the application of *flaA* genotyping to isolates from a case-control study may allow detection of further sources through the examination of genotype-specific risk factors. Study isolates were collected prospectively from sporadic cases recruited into a case-control study and genotyped using *flaA* restriction fragment-length polymorphism (RFLP; *flaA* genotyping) analysis, a moderate throughput, low cost method [29] with good correlation with MLST [30]. Genotyping data were linked to exposure data for the investigation of risk factors.

METHODS

Study base

Data and *Campylobacter* isolates collected for this study were drawn from an Australian case-control study that was conducted in five states between September 2001 and September 2002. A description of the study and risk factors for subjects aged ≥5 years was reported by Stafford *et al.* [13]. The study recruited cases and controls from five of the eight states

and territories in Australia. The largest state, New South Wales, was excluded because campylobacteriosis is not notifiable there and the Northern Territory and Australian Capital Territory were not included in the study because too few cases are notified. As previously described [13], each site aimed to prospectively recruit 200 cases of all ages using systematic sampling from a notifiable disease register; every second case from Tasmania and Victoria, every fourth and sixth case from Queensland and South Australia, respectively and every case from Western Australia. The sampling strategy was based on the expected number of cases notified from participating laboratories and the number required to detect significant associations with hypothesized risk factors [13]. Furthermore, sample size calculations determined that about 550 isolates would probably be collected and that number would prove sufficient to detect significant associations between some genotypes and risk factors.

Cases were defined as individuals with diarrhoea (≥3 or more loose stools in a 24 h period), who had culture-confirmed *C. jejuni* infection (and no other pathogen such as *Salmonella*, *Shigella*, or the enteric protozoa detected), whose isolates were *flaA* genotyped, whose stool samples were collected within 10 days of diarrhoea onset and who were interviewed within 30 days of onset. Telephone interviews were conducted when verbal consent was given by the study subject (or carer if the case was aged <16 years).

Controls were drawn (about one per case) from a control bank generated during a national cross-sectional survey of gastroenteritis conducted in 2001; a description of the survey is given by Hall *et al.* [31] and a description of control selection is reported by Stafford *et al.* [13]. In brief, households were selected using random digit dialling during the gastroenteritis survey, a household member was asked to participate in the survey and subsequently consent to be part of a control bank. From this control bank, potential controls were selected and frequency-matched to cases by age bands (0–4, 5–9, 10–19, 20–29, 30–59 and ≥60 years) in each state at a ratio of 1:1. If a person did not wish to participate or was excluded on the criteria (described below), a subsequent person was sought from the control bank. Once a control had been selected from a household, that household was no longer eligible for future selection of controls. Controls were interviewed within 30 days of interview of a notified case.

Cases and controls were excluded if they did not have a phone number, were unable to be contacted

after at least six attempts, they/their parents were non-English-speakers, they could not answer questions, or if a household member had had diarrhoea or a confirmed *Campylobacter* infection in the 4 weeks prior to onset. In addition, cases were excluded if they were unable to recall the diarrhoea onset date or they were part of an outbreak.

Questionnaire

A standard questionnaire was administered by telephone to obtain information on a range of variables, including host factors (underlying illnesses, prior consumption of antimicrobial agents, antacids and immune suppressive therapies), overseas travel, dining locations, consumption of water and food (fruit and vegetables, meat, poultry, seafood, eggs and dairy products), animal and pet exposures, and demographics. Cases were asked additional questions about their illness and treatment. For cases, all questions related to the 7-day period before their onset of diarrhoea, except for prior use of antibiotics, antacids or immune suppressive therapies, which were based on the preceding 4 weeks. Exposure information was not collected from subjects who had travelled outside Australia during the 7-day exposure period. Controls were asked about the 7 days or 4 weeks prior to interview.

Laboratory methods

Isolates

Isolates from diarrhoeal stool cultures were stored and subsequently identified to species level by hippurate hydrolysis and PCR as described previously [32, 33]. Those identified as *C. jejuni* using these methods were *flaA* genotyped.

flaA genotyping

Flagellin A RFLP typing (*flaA* genotyping) was performed according to the method described by Nachamkin *et al.* [15]. Briefly, this involved PCR amplification of *flaA*, followed by digestion of products with the restriction enzyme *Dde*I and separation of fragments by agarose gel electrophoresis. Genotyping was performed by five laboratories, using standardized reagents and methods (including DNA extraction, PCR, restriction enzyme digestion, agarose gel electrophoresis and photography), which had been optimized by one laboratory.

Quality assurance

A set of eight isolates was distributed to the laboratories that genotyped isolates. Tiff images of *flaA* genotyping gels were loaded onto a BioNumerics database located at one laboratory and examined for comparability. Feedback was provided on accuracy of the patterns and image quality. Reaction or photography conditions were modified, if necessary, to produce images consistent with those from prior testing of the quality assurance set. The positive control, NCTC 11351, was included in each test run. GeneRuler™ 100 bp DNA LadderPlus (MBI Fermentas, Vilnius, Lithuania) was included in lanes 1, 5, 10, and 15 of all gels. If the gel pattern of the positive control was not compatible with those included in the BioNumerics database, gel images were to be rejected, however, this did not occur.

Assignment of flaA genotypes

Tiff images of gels were loaded onto a BioNumerics database and patterns were normalized according to molecular-weight standards on each gel. *flaA* genotypes were designated by a number that was assigned arbitrarily; numbered genotypes differed from each other by at least two bands and subtypes of numbered genotypes, designated by a letter (e.g. *flaA*-6 and *flaA*-6b) differed by a single band only. The numbering of genotypes was consistent with a previous Australian study [29, 30] and isolates from the previous study were included in the BioNumerics library used for analysis in this study. Genotypes were grouped together using the Dice band matching coefficient and UPGMA clustering method with a position tolerance of 1% and an optimization of 1% which clustered at >90% similarity [29]. Resultant dendrograms were checked visually by two researchers and about 10% were re-verified by the second researcher in the case of discrepancies.

Statistical analyses

Logistic regression analysis was used to compare demographic characteristics and host factors between study cases and (*a*) study controls, (*b*) cases that did not have isolates *flaA* genotyped and (*c*) cases notified through the national surveillance system.

Logistic regression analysis was also used to identify potential risk factors for specific *flaA* genotypes. The exposures reported for cases of each major *flaA* genotype were compared to those for all study

controls combined, in order to increase the power of hypothesis testing. A final model for each genotype was constructed by including all exposure variables with $P < 0.1$ (in univariate analyses) and using backwards stepwise elimination, controlling for confounders (demographic and host factors). Models were tested for goodness of fit and compared using the likelihood ratio test. Genotype-specific population-attributable fractions (PAF) were calculated for each risk factor from final models for each of the major *flaA* genotypes and for the group comprising 'other' *flaA* genotypes.

To allow for the possibility that some food and environmental exposures were location-specific, we included terms for the interaction between exposure and state (as a categorical variable) in the logistic regression models. The significance of multiple interaction terms was tested using the likelihood ratio test. Only significant interactions are reported.

Multinomial regression was applied to data on cases to explore differences in exposure variables for the major *flaA* genotypes, using 'other' *flaA* genotypes, comprising the remaining study cases, as the reference group, controlling for confounders. This type of analysis was also used to compare demographic and clinical characteristics between cases infected with each of the major *flaA* genotypes. Results are expressed as relative risk ratios (RRRs), as is appropriate for this case-only analysis [34].

Analyses were performed using Stata version 9.1 (Stata Corporation, College Station, TX, USA).

RESULTS

Recruitment of cases and controls

During the study period there were 8847 *Campylobacter* notifications in the five participating states and, of these, 1019 cases were recruited and interviewed (12%). There were 590 (58%) cases for which an isolate was stored, subsequently found to be *C. jejuni*, and *flaA* genotyped, representing 7% of notifications. There were no significant differences in age distribution between 590 study cases and cases notified in 2001 but fewer males were included (52·5% among cases *vs.* 54·4% for notifications, $P = 0.03$). When cases for whom isolates were genotyped were compared to cases that did not have an isolates genotyped for demographic, host factor and clinical characteristics, a greater number of the

former were educated to school level only [37% *vs.* 28%; odds ratio (OR) 1·5, 95% confidence interval (CI) 1·1–2·0] but no further differences were detected.

A total of 967 controls were recruited and there were statistically significant, but modest, differences between study cases and controls with respect to sex, income, place of residence and use of acid-reducing medications (Table 1).

flaA genotype distribution

Among 590 isolates, there were 61 different *flaA* genotypes, of which five accounted for 325 (55%; an image of the electrophoretic patterns of the major genotypes is shown in the Fig.), and 21 (4%) were single isolates. The five major *flaA* genotypes comprised *flaA*-6b ($n = 129$, 22%), *flaA*-6 ($n = 70$, 12%), *flaA*-10 ($n = 48$, 8%), *flaA*-2 ($n = 43$, 7%), *flaA*-131 ($n = 35$, 6%); the remaining 265 study cases comprised the 'other' genotype group used in case-only analyses. Some geographic differences were noted for the major genotypes; *flaA*-10, and -131 were identified in all states, *flaA*-6b and -2 were found in four of the five states but *flaA*-6b, the most common genotype, was not found in South Australia and *flaA*-6 was found in South Australia and Queensland only. Of the major *flaA* genotypes, only *flaA*-2, and -10 were detected among the 13 isolates from overseas travellers (data not shown). The full descriptive epidemiology of *flaA* genotypes and detection of clusters will be reported separately.

The five major *flaA* genotypes (*flaA*-2, -6, -6b, -10, and -131) were analysed separately for risk factors.

Patient characteristics and symptom profile for the major *flaA* genotypes

Comparison of cases due to each of the five major *flaA* genotypes with those due to all 'other' genotypes using multinomial regression showed no differences in the following characteristics (proportions for all study cases are shown in parentheses): proportion of males (53%), those with cramps (89%), persistent diarrhoea (12%), those that were hospitalized (13%), those treated with anti-diarrhoeal medications (48%) or intravenous fluids (12%). A significantly higher proportion of subjects with *flaA*-2, compared to 'other' genotypes, had fever (RRR 2·3, 95% CI 1·1–5·4), bloody diarrhoea (RRR 2·2, 95%

Table 1. *Comparison of characteristics of study cases with controls*

Characteristic	Study cases		Controls		OR	95% CI
	No.	%	No.	%		
Male	310/590	52·5	455/967	47·1	**1·2**	1·0–1·5
Age group (yr)						
0–4	80/590	13·6	129/967	13·3	Ref.	
5–9	35/590	5·9	69/967	7·1	0·8	0·5–1·3
10–19	56/590	9·5	95/967	9·8	0·9	0·6–1·5
20–29	102/590	17·3	144/967	14·9	1·1	0·8–1·7
30–59	226/590	38·3	384/967	39·7	0·9	0·7–1·3
⩾60	91/590	15·4	146/967	15·1	1·0	0·7–1·5
Aboriginal/Torres Strait Islander	15/579	2·6	26/958	2·7	1·0	0·5–1·9
English only spoken at home	516/579	89·1	865/959	90·2	0·9	0·6–1·3
Education						
Post-secondary school	360/577	62·4	591/950	62·2	1·0	0·8–1·2
Residence						
City/urban	36/581	6·2	72/961	7·5	Ref.	
Suburban	329/581	56·6	571/961	59·4	1·2	0·8–1·8
Town	134/581	23·1	160/961	16·7	**1·7**	**1·1–2·7**
Rural or remote community	45/581	7·8	115/961	12·0	0·8	0·5–1·3
Farm	37/581	6·4	43/961	4·5	1·7	0·9–3·1
Income, AUS$ (2000/2001)						
<$25 000	103/495	20·8	204/826	24·7	Ref.	
$25 000–50 000	139/495	28·1	263/826	31·8	1·0	0·9–1·4
$51 000–100 000	192/495	38·8	267/826	32·3	**1·4**	**1·1–1·9**
>$100 000	61/495	12·3	92/826	11·1	1·3	0·9–2·0
Host factors						
Diabetes	28/456	6·1	49/663	7·4	0·8	0·5–1·4
Immunosuppressive condition	169/464	36·4	238/673	35·4	1·0	0·8–1·3
Prior antibiotics	32/582	5·5	58/967	6·0	0·9	0·6–1·4
Acid-reducing medication	99/499	19·8	100/828	12·1	**1·8**	**1·3–2·5**

OR, Odds ratio (shown in bold where $P < 0.05$); CI, confidence interval; Ref., reference category.

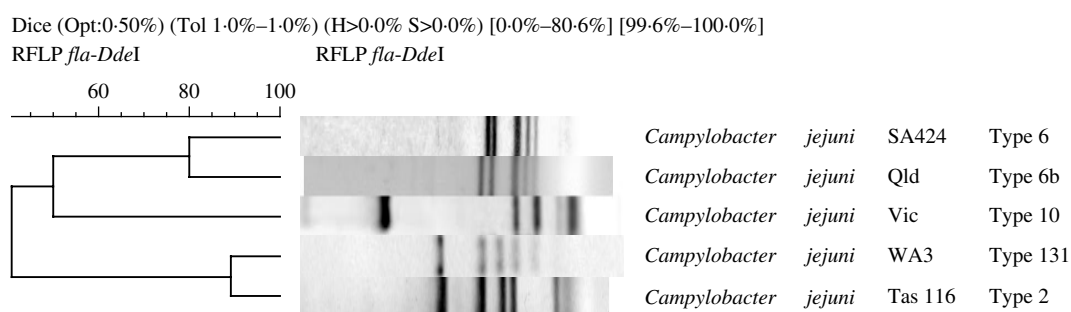Dice (Opt:0·50%) (Tol 1·0%–1·0%) (H>0·0% S>0·0%) [0·0%–80·6%] [99·6%–100·0%]

Fig. *flaA* genotype profiles of the five major *flaA* genotypes.

CI 1·1–4·4) and >20 bowel motions in a 24 h period (RRR 5·4, 95% CI 1·1–25·7).

Sources of *C. jejuni* and risk factors

Subjects that had travelled internationally were included in the study but exposure information was not collected. When compared to controls, overseas travel was significantly associated with *flaA*-10 disease (OR 14·5, 95% CI 2·3–85·7). A total of 66 exposure variables were examined in univariate analyses of cases infected with the major *flaA* genotypes who acquired their infections locally, compared to controls.

Table 2. *Final multiple logistic regression models* for exposures associated with major* Campylobacter jejuni flaA *genotypes*

Exposure‡	OR and 95% CI for exposures among those infected with *C. jejuni flaA* genotypes [Population-attributable fraction, 95% CI]†					
	flaA-2§ (n=43)	*flaA*-6 (n=70)	*flaA*-6b‖ (n=129)	*flaA*-10¶ (n=48)	*flaA*-131# (n=35)	Other** (n=265)
Chicken		3·6 (1·3–10·3) **[65·5%, 11·6–86·5]**				
Undercooked chicken	4·7 (1·7–13·4) **[17·0%, 2·2–30·9]**				3·6 (1·1–12·9) **[9·8%, n.c.]**	3·6 (1·9–6·6) **[7·7%, 3·3–11·7]**
Barbecued chicken			1·9 (1·2–2·9) **[16·8%, 3·5–28·3]**			
Undercooked beef				2·6 (1·2–5·5) **[17·2%, 0·5–31·2]**		
Diced beef						0·6 (0·3–0·8)
Any sausage		0·3 (0·1–0·7)				
Veal			0·3 (0·1–0·9)			
Offal			2·9 (1·1–7·9) **[4·5%, n.c.]**	11·4 (4·1–31·8) **[15·9%, 5·5–25·2]**		2·5 (1·1–6·0) **[2·7%, n.c.]**
Paté			2·2 (1·1–4·6) **[3·8%, n.c.]**	3·3 (1·2–9·6) **[7·8%, n.c.]**		
Ham	2·4 (1·1–5·2) **[39·3%, 1·0–62·8]**					
Organic fruit/vegetables			0·4 (0·2–0·9)			0·5 (0·3–0·9)
Bottled water						2·5 (1·5–4·3) **[6·0%, 2·0–10·0]**
Puppy	4·0 (1·1–14·8) **[9·0%, n.c.]**					
Pet chicken				2·1 (1·1–4·3) **[5·3%, n.c.]**		
Pet chicken <6 months				20·6 (3·0–117·4) **[8·2%, 0·4–15·4]**	10·2 (2·6–40·4) **[13·4%, 0·8–24·4]**	
Farm birds		11·3 (2·5–51·9) **[5·8%, n.c.]**				
Poor barbeque technique	4·4 (1·2–16·3) **[8·7%, n.c.]**			4·6 (1·2–18·4) **[7·6%, n.c.]**	4·4 (1·1–17·5) **[6·9%, n.c.]**	

OR, Odds ratio, CI, confidence interval, n.c., 95% CI unable to be calculated due to small numbers.

* Logistic regression models controlled for age, sex (and state, when appropriate).

† Population-attributable fraction shown in bold within square brackets with 95% CI, determined for cases infected with each *flaA* genotype; blank cells indicate that no statistically significant association was found for the respective *flaA* genotype and exposure.

‡ Exposure period of 7 days.

§ Model for this *flaA* genotype included state.

‖ Model for this *flaA* genotype included consumption of acid-reducing medications.

¶ Model for this *flaA* genotype included chronic gastrointestinal illness and immunosuppressive therapies.

Model for this *flaA* genotype included age and place of residence.

** All remaining *flaA* genotyped cases. Model for this *flaA* genotype included chronic gastrointestinal illness.

In final multivariate models, constructed to explain exposures associated with locally acquired disease (Table 2), infection with *flaA*-6b, the most common *flaA* genotype, was independently associated with consumption of barbecued chicken, offal, paté and exposure to pet chickens. The second most common *flaA* genotype, *flaA*-6, was independently associated with consuming chicken and exposure to farm birds (ducks, geese, etc.). Infection with *flaA*-10 was associated with undercooked beef, offal, paté, exposure to young pet chickens (aged <6 months), and poor food handling. *flaA*-131 was associated with chicken (meat and pets) and poor food handling, and *flaA*-2 was associated with consumption of undercooked chicken, ham, exposure to puppies and poor food-handling practices (Table 2). When all 'other' *flaA* genotypes combined (comprising 56 *flaA* genotypes) were compared to controls, disease among this group was associated with undercooked chicken, offal, and bottled water (Table 2). Of the seven *flaA*-10 case-patients that consumed offal, three specified lamb, one chicken and the remaining three cases did not indicate a species. Of the seven *flaA*-6b case-patients that consumed offal, two specified lamb, the remaining gave no details on species. Among *flaA*-6b and -10 cases that ate offal, there was no geographic or time clustering. While the odds ratios for some of the risk factors implicated may be of borderline statistical significance, some (e.g. offal and contact with pet chickens and farm birds) were unusually high. In the final multivariate models for each of the major *flaA* genotypes, 74% of *flaA*-2 cases, 71% of *flaA*-6 cases, 30% of *flaA*-6b cases, 57% of *flaA*-10 cases, 30% of *flaA*-131 cases, and 16% of 'other' *flaA* genotypes were attributed to significant exposures (Table 2).

For the case-only comparisons, we used multinomial regression analyses to determine distinct exposures among the major *flaA* genotypes, as shown in Table 3. *flaA*-2, *flaA*-6b and *flaA*-10 differed from the comparison group in their exposure to various types of non-poultry meats. Significant exposures associated with *flaA*-10 were different from those associated with *flaA*-6b and *flaA*-131, and reflected exposures associated with disease due to those *flaA* genotypes in the final models (Table 2). *flaA*-2 infected cases were significantly more likely to consume ham (Tables 2 and 3). Poultry meat exposures did not differ between *flaA* types in multinomial regression analyses (Table 3), but were significantly associated with disease due to *flaA*-2, -6, -6b and -131 genotypes when compared to healthy controls (Table 2).

DISCUSSION

Case-control studies of campylobacteriosis have consistently identified chicken consumption as the most commonly associated risk factor [7, 9, 11, 13, 35]. Similarly we found chicken associated with disease due to four out of five major *flaA* genotypes, suggesting that chicken harbour a range of *C. jejuni* genotypes. Interventions aimed at minimizing chicken contamination are needed to reduce the burden of disease, and have been initiated in Iceland and Denmark [36, 37]. Here we have attempted to determine whether separate analysis of *C. jejuni* genotypes for risk factors may provide insights into further sources of this important disease.

We demonstrated the benefit of using molecular methods to more specifically define cases of campylobacteriosis to study possible risk factors for infection. We found that illness due to *C. jejuni* genotype *flaA*-10 was independently associated with undercooked beef consumption. In both case-control and case-only comparisons undercooked beef was associated with *flaA*-10 disease (Tables 2 and 3). Non-poultry meat has not previously been identified as a risk factor for campylobacteriosis in Australia [13] (L. E. Unicomb *et al.*, unpublished results). Consumption of raw milk and/or contact with calves have been implicated in four Australian outbreaks [38] (OzFoodNet Outbreak Register, M. Kirk, personal communication, July 2006) and exposure to non-poultry meats and bovine husbandry have been associated with *Campylobacter* illness in case-control studies conducted in other countries [5, 8, 9, 11]. By way of comparison, the case-control study from which subjects in this study were drawn had 881 cases and 833 controls aged >5 years. It found undercooked chicken, offal, ownership of domestic chickens aged <6 months, and domestic dogs aged <6 months as risk factors [13].

Disease caused by *flaA*-2 was associated with exposure to ham in both case-control and case-only comparisons (Tables 2 and 3). This finding was unexpected, since pigs are predominantly (but not exclusively) infected with *C. coli* [39]. Previously, a ham-containing salad has been implicated in a *C. coli* outbreak [40] and it has been detected as a risk factor in a case-control study conducted in the United States [9]. It is unclear how processed meats such as ham could be the source of disease; contamination at retail outlets from raw meats may occur. Alternatively, this finding may reflect cross-contamination

Table 3. *Multinomial logistic regression analysis* for the major* Campylobacter jejuni flaA *genotypes compared to 'other'* flaA *genotypes*

Exposure	flaA-2 (n=43) Cases exposed	flaA-2 (n=43) RRR‡ (95% CI)	flaA-6 (n=70) Cases exposed	flaA-6 (n=70) RRR (95% CI)	flaA-6b (n=129) Cases exposed	flaA-6b (n=129) RRR (95% CI)	flaA-10 (n=48) Cases exposed	flaA-10 (n=48) RRR (95% CI)	flaA-131 (n=35) Cases exposed	flaA-131 (n=35) RRR (95% CI)	Other† (n=265) Cases exposed
Poultry meat											
Chicken	90%	1·5 (0·5–4·5)	93%	2·4 (0·8–7·2)	82%	0·8 (0·4–1·4)	85%	1·0 (0·4–2·5)	85%	0·9 (0·3–2·8)	85%
Undercooked	12%	1·0 (0·4–2·8)	6%	0·5 (0·2–1·5)	6%	0·5 (0·2–1·2)	5%	0·4 (0·1–1·5)	9%	0·8 (0·2–2·9)	12%
Barbecued	38%	0·6 (0·3–1·3)	37%	0·6 (0·3–1·0)	47%	0·9 (0·6–1·4)	46%	1·0 (0·5–1·9)	51%	0·9 (0·4–1·8)	49%
Non-poultry meat											
Undercooked beef	17%	1·1 (0·5–2·7)	14%	0·9 (0·4–2·1)	30%	1·1 (0·6–2·0)	21%	**2·5 (1·2–5·2)**	16%	1·7 (0·7–4·3)	15%
Diced beef	21%	1·7 (0·7–4·2)	7%	0·5 (0·2–1·5)	22%	1·8 (1·0–3·4)	11%	0·8 (0·3–2·5)	26%	2·2 (0·8–5·7)	13%
Any sausage	34%	1·0 (0·5–2·0)	14%	**0·3 (0·1–0·6)**	35%	1·0 (0·6–1·6)	13%	**0·3 (0·1–0·7)**	45%	1·4 (0·5–3·0)	35%
Veal	5%	0·7 (0·2–3·3)	12%	1·8 (0·7–4·5)	4%	0·5 (0·2–1·5)	11%	1·7 (0·5–5·3)	4%	0·5 (0·1–3·8)	7%
Offal	0%	—	3%	0·6 (0·1–2·9)	5%	1·2 (0·5–3·2)	16%	**3·9 (1·4–10·8)**	0%	—	5%
Paté	5%	1·1 (0·2–5·3)	6%	1·4 (0·4–4·6)	9%	**2·3 (1·1–5·5)**	13%	**3·4 (1·2–9·8)**	0%	—	4%
Ham	73%	**2·6 (1·3–5·5)**	5%	1·0 (0·6–1·7)	53%	1·1 (0·7–1·7)	60%	1·5 (0·8–2·9)	45%	0·8 (0·4–1·6)	51%
Other food											
Organic fruit/ vegetables	10%	1·6 (0·5–5·0)	7%	1·1 (0·4–3·3)	6%	0·8 (0·3–2·1)	10%	1·5 (0·5–4·7)	6%	0·9 (0·2–4·2)	7%
Environmental											
Bottled water	12%	1·2 (0·4–3·3)	11%	1·1 (0·5–2·6)	6%	0·6 (0·2–1·3)	16%	1·5 (0·6–3·7)	9%	0·9 (0·3–3·2)	6%
Puppy	8%	0·9 (0·2–2·4)	5%	0·5 (0·2–1·9)	4%	0·4 (0·2–1·2)	7%	0·7 (0·2–2·7)	3%	0·3 (0·0–2·3)	8%
Pet chicken	7%	0·7 (0·2–2·4)	7%	0·7 (0·2–1·8)	11%	1·1 (0·6–2·2)	9%	0·9 (0·3–2·7)	20%	1·9 (0·8–4·9)	10%
Pet chicken <6 months	0%	—	4%	0·5 (0·3–7·3)	2%	0·7 (0·2–2·8)	4%	1·5 (0·3–7·3)	11%	3·1 (0·8–11·2)	3%
Farm birds	2%	1·6 (0·2–14·8)	4%	2·9 (0·6–13·2)	1%	0·5 (0·1–4·6)	2%	1·5 (0·2–13·4)	0%	—	2%
Poor barbeque technique	9%	2·0 (0·6–6·5)	1%	0·3 (0·0–2·2)	3%	0·6 (0·2–2·0)	6%	1·3 (0·4–4·8)	9%	2·2 (0·6–8·5)	5%

* Multinomial models controlled for age and sex.
† All cases with *flaA* genotypes apart from the five major *flaA* genotypes.
‡ Relative risk ratio (RRR) with 95% confidence intervals (CI) given in parentheses. RRR and 95% CIs given in bold indicate that a statistically significant difference was found for the respective *flaA* genotype when compared to 'other' *flaA* genotypes.

of ham and other foods during preparation in the home.

Consumption of offal (from a variety of species) was associated with disease due to *flaA*-6b, -10 and 'other' genotypes, and was also reported for all *Campylobacter* species and genotypes in an Australian case-control study [13]. Furthermore, duck liver consumption has been implicated in one Australian outbreak (OzFoodNet Outbreak Register, M. Kirk, personal communication, July 2007).

Gel patterns for *flaA*-6 and *flaA*-6b genotypes differed by the size of one band (Fig.), probably from an insertion or deletion in the *flaA* gene. *flaA*-6 was detected in a previous study from January 1999 to July 2001 in New South Wales [29, 30] and has been detected in a subsequent study in South Australia from November 2005 to March 2006 (B. Coombs, personal communication, December 2006). Geographic and temporal differences in *flaA*-6 and -6b distribution suggest that one may be a variant of the other. When *flaA*-6 and -6b were analysed as a single genotype (controlling for the potential confounders age and sex) similar variables were significantly associated with disease, in univariate models, as those for the more common *flaA*-6b alone (data not shown). The two genotypes were analysed separately as we could not assume that they are variants without further laboratory investigations.

Clinical manifestations differed slightly among those infected with the major *flaA* genotypes with *flaA*-2 infections apparently more severe than those due to other genotypes; this suggests that there may be differences in virulence between *flaA* genotypes. While the differences between *flaA*-2 study cases and others were small, they were consistent for symptoms that resulted in missing school, work, recreation or other activities. We have previously reported that *flaA* genotypes closely predicted MLST; 88% of *flaA*-2 isolates, when tested by MLST, belonged to clonal complex (CC) 48 [30]. Among CC 48 human isolates on the international MLST database (http://pubmlst.org/campylobacter; accessed 11 July, 2007), those included had caused Guillain–Barré syndrome, Miller–Fisher syndrome and systemic disease in addition to gastroenteritis. Genotype-specific differences in symptoms by age could not be explored in this study as numbers in each age group were small.

Our findings should be considered in the light of study limitations. Selection bias in recruitment of controls is possible, since people who spend more time at home would be more easily contactable, but several factors were controlled for in analyses. Measurement biases may have occurred as we relied on information from participants that was not validated. However, this applied to both cases and controls. Interviewer bias may have occurred as interviewers knew which interviewees were cases and controls; and recall bias was possible since cases potentially would have better recall than controls. Study cases were selected from notified cases, which include those with more severe disease. Study cases for whom isolates were not genotyped, had similar characteristics to those that were genotyped, a greater number of the former were educated to school level only; this small difference may have had an impact on exposures. We conducted many hypothesis tests, so it is plausible that some statistically significant differences may have arisen by chance; of the 66 exposure variables examined for each subtype, an average of three are expected due to chance alone (using the 5% level of significance). Our findings therefore need to be replicated by further studies to confirm our results.

Detecting genotype-specific risk factors may be better determined by enrolling cases based on their isolate subtype results, focusing on the more common subtypes, thereby including sufficient sample size to test hypotheses. This would further be enhanced by studies of the distribution of *flaA* genotypes of non-human *C. jejuni* isolates; such data on Australian non-human isolates are not available. Information on genotypes for other countries is available from the international MLST database. As stated above, *flaA* genotypes can closely predict MLST CCs: 96% of Australian *flaA*-6 isolates belong to CC 257, 91% of *flaA*-10 to CC 354 and 88% of *flaA*-2 to CC 48 [30]. Among the data from various countries on the MLST database, these clonal complexes have been detected in a variety of non-human samples and countries: CC 257 (*flaA*-6) was detected from bovine, avian (poultry and other avian), ovine and porcine samples, CC 354 (*flaA*-10) from bovine, poultry (including environmental) and ovine samples and CC 48 (*flaA*-2) from bovine, avian (poultry, poultry environment and other avian), ovine, water, sand and domestic pet samples. While these data could potentially be useful in supporting findings from the case-control and case-case analyses, only one of the 124 non-human isolates from the CC 257, 354 and 48 isolates on the database was from Australia. Determining *flaA* genotypes from systematically collected, non-human sources in Australia may be useful to identify potential reservoirs of genotypes.

flaA genotyping is a gel-based method which has limitations; it requires standardization to achieve comparable results across laboratories but is rapid and cheap. Ideally, this method would be automated to achieve 'high throughput' status. There are reports of instability in the *flaA* locus [41] but the majority of *C. jejuni* isolates are apparently genetically stable in this region over time [26, 42, 43] and that instability may be strain specific. Sequence-based methods such as MLST remove dependence on visual and potentially subjective, genotype assignment but require expensive equipment to handle medium to high numbers of isolates and reagents are costly [29, 44]. Isolates from this study were genotyped retrospectively, but 'real time', genotyping of notified case isolates would be preferable, to enable rapid detection of temporal clusters using a library of common *flaA* genotypes and should be feasible.

This is the first study to suggest the value of *flaA* genotyping for identifying strain-specific risk factors for *Campylobacter*. Case-control analyses using logistic regression and case-only analyses using multinomial regression were employed to detect risk factors for *C. jejuni flaA* genotypes among a selection of notified cases drawn from five Australian states over a 12-month period. Differences were detected for symptom profile, geographic distribution and exposures between *flaA* genotypes. The value of *flaA* genotyping is therefore worthy of further investigations in studies with a larger sample size and in other settings, and particularly in the course of outbreak investigations. The ability of *flaA* genotyping to detect clusters and outbreaks among apparently sporadic notified cases will be assessed in a separate report (L. E. Unicomb *et al.*, unpublished observations).

APPENDIX. Australian Campylobacter Subtyping Study Group (*listed in alphabetical order*):

Penny Adamson (Flinders Medical Centre, South Australia), Rosie Ashbolt (Public and Environmental Health Service, Department of Health and Human Services, Hobart), Kellie Cheung (Institute of Clinical Pathology and Medical Research, Westmead, New South Wales), Barry Combs (Department of Human Services, Adelaide, South Australia), Craig Dalton (Hunter New England Population Health, Newcastle, New South Wales), Steve Djordjevic (Elizabeth Macarthur Agricultural Institute, Camden, New South Wales), Robyn Doyle (Institute of Medical and Veterinary Science, Adelaide, South Australia), John Ferguson (Hunter New England Health Service, Newcastle, New South Wales), Lyn Gilbert (Institute of Clinical Pathology and Medical Research, Westmead, New South Wales), Rod Givney (Department of Human Services, Adelaide, South Australia), David Gordon (Flinders Medical Centre, Bedford Park, South Australia), Joy Gregory (Department of Human Services, Melbourne, Victoria), Geoff Hogg (Microbiological Diagnostic Unit, University of Melbourne, Parkville, Victoria), Tim Inglis (Division of Microbiology & Infectious Diseases, PathWest, Nedlands, Western Australia), Peter Jelfs (Institute of Clinical Pathology and Medical Research, Westmead, New South Wales), Martyn Kirk (Department of Health and Ageing, Canberra, Australian Capital Territory), Karin Lalor (Department of Human Services, Melbourne, Victoria), Jan Lanser (Institute of Clinical Pathology and Medical Research, Westmead, New South Wales), Lance Mickan (Institute of Medical and Veterinary Science, Adelaide, South Australia), Lyn O'Reilly (Division of Microbiology & Infectious Diseases, PathWest, Nedlands, Western Australia), Rosa Rios (Microbiological Diagnostic Unit, Parkville, Victoria), Minda Sarna (Department of Health, Perth, Western Australia), Hemant Sharma (Hunter New England Health Service, Newcastle New South Wales), Helen Smith (Queensland Health Scientific Services, Coopers Plains, Queensland), Leanne Unicomb (Hunter New England Population Health and National Centre for Epidemiology and Population Health, Australian National University, Canberra, ACT), Mary Valcanis (Microbiological Diagnostic Unit, University of Melbourne, Parkville, Victoria).

DECLARATION OF INTEREST

None.

REFERENCES

1. **Owen R, et al.** Australia's notifiable diseases status, 2005: annual report of the National Notifiable Diseases Surveillance System. *Communicable Diseases Intelligence* 2007; **31**: 1–70.
2. **Hall G, et al.** An estimate of under-reporting of foodborne notifiable diseases: Salmonella, Campylobacter and Shiga toxin producing *Escherichia coli* (STEC). Canberra: National Centre for Epidemiology and Population Health, 2006.
3. **Anon.** Communicable Diseases Surveillance: highlights for November, 2000. *Communicable Diseases Intelligence* 2000; **24**: 391–405.
4. **Wingstrand A, et al.** Fresh chicken as main risk factor for campylobacteriosis, Denmark. *Emerging Infectious Diseases* 2006; **12**: 280–285.
5. **Schonberg-Norio D, et al.** Swimming and Campylobacter infections. *Emerging Infectious Diseases* 2004; **10**: 1474–1477.
6. **Michaud S, Menard S, Arbeit RD.** Campylobacteriosis, Eastern Townships, Quebec. *Emerging Infectious Diseases* 2004; **10**: 1844–1847.
7. **Kapperud G, et al.** Factors associated with increased and decreased risk of Campylobacter infection: a prospective case-control study in Norway. *American Journal of Epidemiology* 2003; **158**: 234–242.
8. **Potter RC, Kaneene JB, Hall WN.** Risk factors for sporadic *Campylobacter jejuni* infections in rural Michigan: a prospective case-control study. *American Journal of Public Health* 2003; **93**: 2118–2123.
9. **Friedman CR, et al.** Risk factors for sporadic Campylobacter infection in the United States: A case-control study in FoodNet sites. *Clinical Infectious Diseases* 2004; **38** (Suppl. 3): S285–S296.
10. **Carrique-Mas J, et al.** Risk factors for domestic sporadic campylobacteriosis among young children in Sweden. *Scandinavian Journal of Infectious Diseases* 2005; **37**: 101–110.
11. **Neimann J, et al.** A case-control study of risk factors for sporadic *Campylobacter* infections in Denmark. *Epidemiology and Infection* 2003; **130**: 353–366.
12. **Tenkate TD, Stafford RJ.** Risk factors for *Campylobacter* infection in infants and young children: a matched case-control study. *Epidemiology and Infection* 2001; **127**: 399–404.
13. **Stafford RJ, et al.** A multi-centre prospective case-control study of *Campylobacter* infection in persons aged 5 years and older in Australia. *Epidemiology and Infection* 2007; **135**: 978–988.
14. **Rodrigues LC, et al.** The study of infectious intestinal disease in England: risk factors for cases of infectious intestinal disease with *Campylobacter jejuni* infection. *Epidemiology and Infection* 2001; **127**: 185–193.
15. **Nachamkin I, Bohachick K, Patton CM.** Flagellin gene typing of *Campylobacter jejuni* by restriction fragment length polymorphism analysis. *Journal of Clinical Microbiology* 1993; **31**: 1531–1536.
16. **Dingle KE, et al.** Multilocus sequence typing system for *Campylobacter jejuni*. *Journal of Clinical Microbiology* 2001; **39**: 14–23.
17. **Hanninen ML, et al.** A three-year study of *Campylobacter jejuni* genotypes in humans with domestically acquired infections and in chicken samples from the Helsinki area. *Journal of Clinical Microbiology* 2000; **38**: 1998–2000.
18. **Clark CG, et al.** Use of the Oxford multilocus sequence typing protocol and sequencing of the flagellin short variable region to characterize isolates from a large outbreak of waterborne *Campylobacter* sp. strains in Walkerton, Ontario, Canada. *Journal of Clinical Microbiology* 2005; **43**: 2080–2091.
19. **Hedberg CW, et al.** Limitations of pulsed-field gel electrophoresis for the routine surveillance of Campylobacter infections. *Journal of Infectious Diseases* 2001; **184**: 242–244.
20. **de Boer P, et al.** Computer-assisted analysis and epidemiological value of genotyping methods for *Campylobacter jejuni* and *Campylobacter coli*. *Journal of Clinical Microbiology* 2000; **38**: 1940–1946.
21. **Lindstedt BA, et al.** Comparative fingerprinting analysis of *Campylobacter jejuni* subsp. *jejuni* strains by amplified-fragment length polymorphism genotyping. *Journal of Clinical Microbiology* 2000; **38**: 3379–3387.
22. **Frost JA, Gillespie IA, O'Brien SJ.** Public health implications of *Campylobacter* outbreaks in England and Wales, 1995–1999: epidemiological and microbiological investigations. *Epidemiology and Infection* 2002; **128**: 111–118.
23. **Sails AD, Swaminathan B, Fields PI.** Utility of multilocus sequence typing as an epidemiological tool for investigation of outbreaks of gastroenteritis caused by *Campylobacter jejuni*. *Journal of Clinical Microbiology* 2003; **41**: 4733–4739.
24. **Colles FM, et al.** Genetic diversity of *Campylobacter jejuni* isolates from farm animals and the farm environment. *Applied and Environmental Microbiology* 2003; **69**: 7409–7413.
25. **Schouls LM, et al.** Comparative genotyping of *Campylobacter jejuni* by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination. *Journal of Clinical Microbiology* 2003; **41**: 15–26.
26. **Fitzgerald C, et al.** Use of pulsed-field gel electrophoresis and flagellin gene typing in identifying clonal groups of *Campylobacter jejuni* and *Campylobacter coli* in farm and clinical environments. *Applied and Environmental Microbiology* 2001; **67**: 1429–1436.
27. **Nielsen EM, et al.** Most *Campylobacter* subtypes from sporadic infections can be found in retail poultry

products and food animals. *Epidemiology and Infection* 2006; **134**: 758–767.

28. **Manning G, *et al.*** Multilocus sequence typing for comparison of veterinary and human isolates of *Campylobacter jejuni*. *Applied and Environmental Microbiology* 2003; **69**: 6370–6379.

29. **O'Reilly LC, Inglis TJ, Unicomb L.** Australian multicentre comparison of subtyping methods for the investigation of *Campylobacter* infection. *Epidemiology and Infection* 2006; **134**: 768–779.

30. **Djordjevic SP, *et al.*** Clonal complexes of *Campylobacter jejuni* identified by multilocus sequence typing are reliably predicted by restriction fragment length polymorphism analyses of the flaA gene. *Journal of Clinical Microbiology* 2007; **45**: 102–108.

31. **Hall G, *et al.*** Estimating foodborne gastroenteritis, Australia. *Emerging Infectious Diseases* 2005; **11**: 1257–1264.

32. **Burnett TA, *et al.*** Speciating *Campylobacter jejuni* and *Campylobacter coli* isolates from poultry and humans using six PCR-based assays. *FEMS Microbiology Letters* 2002; **216**: 201–209.

33. **Sharma H, *et al.*** Antibiotic resistance in *Campylobacter jejuni* isolated from humans in the Hunter Region, New South Wales. *Communicable Diseases Intelligence* 2003; **27** (Suppl.): S80–S88.

34. **Gould W, Sribney W.** Comparison of the nested logit and the constrained multinomial models. In: *Stata FAQs*. College Station, TX: StataCorp LP, 1999.

35. **Effler P, *et al.*** Sporadic *Campylobacter jejuni* infections in Hawaii: associations with prior antibiotic use and commercially prepared chicken. *Journal of Infectious Diseases* 2001; **183**: 1152–1155.

36. **Stern NJ, *et al.*** *Campylobacter* spp. in Icelandic poultry operations and human disease. *Epidemiology and Infection* 2003; **130**: 23–32.

37. **Krause M, *et al.*** Comparative, collaborative, and on-site validation of a TaqMan PCR method as a tool for certified production of fresh, campylobacter-free chickens. *Applied and Environmental Microbiology* 2006; **72**: 5463–5468.

38. **Dalton CB, *et al.*** Foodborne disease outbreaks in Australia, 1995 to 2000. *Communicable Diseases Intelligence* 2004; **28**: 211–224.

39. **Jensen AN, *et al.*** The occurrence and characterization of *Campylobacter jejuni* and *C. coli* in organic pigs and their outdoor environment. *Veterinary Microbiology* 2006; **116**: 96–105.

40. **Ronveaux O, *et al.*** A *Campylobacter coli* foodborne outbreak in Belgium. *Acta Clinica Belgica* 2000; **55**: 307–311.

41. **Harrington CS, Thomson-Carter FM, Carter PE.** Evidence for recombination in the flagellin locus of *Campylobacter jejuni*: implications for the flagellin gene typing scheme. *Journal of Clinical Microbiology* 1997; **35**: 2386–2392.

42. **Burnens AP, *et al.*** Restriction fragment length polymorphisms among the flagellar genes of the Lior heat-labile serogroup reference strains and field strains of *Campylobacter jejuni* and *C. coli*. *Epidemiology and Infection* 1995; **114**: 423–431.

43. **On SL, *et al.*** Validity of *Sma*I-defined genotypes of *Campylobacter jejuni* examined by *Sal*I, *Kpn*I, and *Bam*HI polymorphisms: evidence of identical clones infecting humans, poultry, and cattle. *Epidemiology and Infection* 1998; **120**: 231–237.

44. **Dingle KE, *et al.*** Molecular characterization of *Campylobacter jejuni* clones: a basis for epidemiologic investigation. *Emerging Infectious Diseases* 2002; **8**: 949–955.

45. **Jolley KA, Chan MS, Maiden MC.** mlstdbNet – distributed multi-locus sequence typing (MLST) databases. *BMC Bioinformatics* 2004; **5**: 86.