JOURNAL OF BACTERIOLOGY, Apr. 2001, p. 2553–2559 Vol. 183, No. 8
0021-9193/01/$04.00+0 DOI: 10.1128/JB.183.8.2553–2559.2001

Allelic Diversity and Recombination in *Campylobacter jejuni*

SEBASTIAN SUERBAUM,[1]* MARC LOHRENGEL,[1] AGNES SONNEVEND,[1]
FLORIAN RUBERG,[2] AND MANFRED KIST[2]

Institut für Hygiene und Mikrobiologie der Universität Würzburg, D-97080 Würzburg,[1] and Institut für Medizinische Mikrobiologie und Hygiene der Universität Freiburg, D-79104 Freiburg,[2] Germany

Received 12 September 2000/Accepted 29 January 2001

The allelic diversity and population structure of *Campylobacter jejuni* were studied by multilocus nucleotide sequence analysis. Sequences from seven housekeeping genes were obtained from 32 *C. jejuni* isolates isolated from enteritis patients in Germany, Hungary, Thailand, and the United States. Also included was strain NCTC 11168, the complete genomic sequence of which has recently been published. For all loci analyzed, multiple strains carried identical alleles. The frequency of synonymous and nonsynonymous sequence polymorphisms was low. The number of unique alleles per locus ranged from 9 to 15. These alleles occurred in 31 different combinations (sequence types), so that all but two pairs of strains could be distinguished from each other. Sequences were analyzed for evidence of recombination by the homoplasy test and split decomposition. These analyses showed that intraspecific recombination is frequent in *C. jejuni* and has generated extensive diversity of allelic profiles from a small number of polymorphic nucleotides.

Campylobacter jejuni infection is one of the most frequent causes of bacterial food-borne diarrheal disease all over the world. While in most cases the disease is self-limiting, *C. jejuni* infections can give rise to the debilitating and potentially fatal Guillain-Barré syndrome, a progressive neuromuscular paralysis (for a review, see reference 32).

Over the last decade, numerous genotypical typing methods for *Campylobacter* species have been described, including pulsed-field gel electrophoresis (12, 35), flagellin gene typing (4, 24), randomly amplified polymorphic DNA (RAPD) PCR analysis (11, 23), and, most recently, amplified fragment length polymorphism (8). (For a review of typing methods for *C. jejuni,* see reference 34.) However, although all of these methods ultimately depend on sequence variation, up to now, there have been very few systematic analyses of nucleotide sequence variability in *C. jejuni*. In several other species of pathogenic bacteria, analysis of nucleotide sequence variation at multiple gene loci has permitted us to gain further understanding of the population structure of these pathogens. Such analyses have shown that different pathogens differ widely in the extent of sequence variation, their population structure, the relative roles of mutation and recombination, and the existence of clonal groupings with distinct geographic distribution patterns (2, 3, 30).

Multilocus sequence typing (MLST), a method that is based on partial nucleotide sequences of multiple (usually around seven) housekeeping genes, has recently been shown to be a powerful technique for bacterial typing (3, 9, 19). Housekeeping genes are preferred over virulence-associated genes, because an analysis of mutations (most of which are usually synonymous, given the strong selection against changes of the amino acid sequence in genes coding for proteins required for growth) in such genes is more likely to adequately reflect the phylogeny of strains. For a more extensive discussion of these arguments, see reference 1. While still relatively expensive, major advantages of this technique are the easy portability of both the method and results and the possibility of building up global databases by using the internet. An additional advantage of MLST approaches is that results can be used to perform phylogenetic and population genetic analyses. The feasibility of sequence-based typing depends on the identification of genes that have sufficiently high sequence variability. Since there was very little information about the extent of sequence variability in *C. jejuni*, we have determined nucleotide sequences of seven housekeeping genes that were selected from the recently published whole genomic sequence of *C. jejuni* (25). Nucleotide sequences of 423- to 660-bp fragments from these housekeeping genes were obtained for a collection of 32 *C. jejuni* strains from Germany, Hungary, Thailand, and the United States and analyzed for their variability. We have also assessed the frequency of recombination with the homoplasy test (21) and studied the population structure of *C. jejuni* by split decomposition. The data show that the population structure of *C. jejuni* is characterized by a low degree of sequence diversity, a relatively small pool of alleles in the housekeeping genes tested, and high rates of intraspecies recombination. Recombination is frequent enough to generate a large number of unique combinations of alleles (sequence types), implying that MLST approaches could be valuable for future studies of the molecular epidemiology of *C. jejuni*.

MATERIALS AND METHODS

Strains. Thirty-two *C. jejuni* strains were used in this study. The sequences of strain NCTC 11168, the complete genomic sequence of which was recently published (25), were added to all data sets. The *C. jejuni* strains used were isolated in two regions of Germany (Würzburg and Freiburg), in Hungary (six strains), in Thailand (five strains), and in the United States. All strains were from patients with enteritis, with the exception of BK612, which was isolated from a blood culture, and SSU9896, a bovine isolate from the United States (24). Available data about the strains are listed in Table 1.

Nucleotide sequencing. Seven fragments of housekeeping genes were selected for the analysis. Details of the fragments sequenced are shown in Table 2. The genes were selected based on the following criteria: they encode housekeeping

* Corresponding author. Mailing address: Institut für Hygiene und Mikrobiologie der Universität Würzburg, Josef-Schneider-Straβe 2, D-97080 Würzburg, Germany. Phone: 49 931 201 3949. Fax: 49 931 201 3445. E-mail: ssuerbaum@hygiene.uni-wuerzburg.de.

TABLE 1. Strains of *C. jejuni* for which sequences were analyzed

Strain	Strain no.	Country of isolation	Penner serotype
95-9783	1	Freiburg, Germany	O7
96-10475	2	Freiburg, Germany	O21
96-11019-2	3	Freiburg, Germany	O1
96-2419	4	Freiburg, Germany	O23
96-3785	5	Freiburg, Germany	ND[a]
96-4260	6	Freiburg, Germany	O7
96-4693	7	Freiburg, Germany	ND
96-4838	8	Freiburg, Germany	O21
96-6530	9	Freiburg, Germany	O14
96-8863-1	10	Freiburg, Germany	O1
96-9092-3	11	Freiburg, Germany	O1
96-9466	12	Freiburg, Germany	O21
97-2541	13	Freiburg, Germany	O37
ATCC 33560	14	United States	O23
BK612	15	Würzburg, Germany[b]	ND
NCTC 11168	16	United Kingdom	O2
CJ99-1464	17	Würzburg, Germany	ND
CJ99-3529	18	Würzburg, Germany	ND
CJ99-3663	19	Würzburg, Germany	ND
CJ99-3773	20	Würzburg, Germany	ND
D 2677[c]	21	United States	O36,23,15w
H107	22	County Pest, Hungary	O3
H110	23	County Pest, Hungary	O3
H50	24	Budapest, Hungary	O2
H56	25	County Pest, Hungary	O1,44
H76	26	County Veszprem, Hungary	O1,44
H98	27	County Pest, Hungary	O2
RV173	28	Thailand	ND
RVB005	29	Thailand	ND
RVB018	30	Thailand	ND
RVB041	31	Thailand	ND
RVB090	32	Thailand	ND
SSU 9896[c]	33	United States[d]	O2

[a] ND, not determined.
[b] Blood culture isolate.
[c] Strain described in reference 24.
[d] Bovine isolate.

TABLE 3. Primers used for amplification and sequencing

Primer	Purpose[a]	Sequence
asd		
CJasd1(+)	A, S	GCA-GGT-GGA-AGT-GTG-AGT-G
CJasd2(−)	A, S	TTT-GTT-GCA-GCA-CCT-ACA-CG
CJasd3(−)	A	ACG-AAT-TTG-ATC-CGC-CAC-AC
CJasd4	A, S	GCC-ATT-GTG-GGT-GCT-ACT-GG
CJasd5	A, S	CGC-TAG-TCA-TTA-AAG-GCA-TAG-G
atpA		
CJatpA1(+)	A, S	GAG-AAG-GTT-TAA-AAG-AAG-GTG-C
CJatpA2(−)	A, S	TGT-AGC-TTT-AAT-TTG-AGC-AGC
ddlA		
CJddla1(+)	A, S	GAT-CAA-TCT-TAT-CCA-TGG-TAG
CJddla2(−)	A, S	AGC-CAA-AGA-ACC-AGG-GTT-TG
eftS (tsf)		
CJefts1(+)	A, S	AAA-GCA-GAT-AGA-CTT-GCT-GC
CJefts2(−)	A, S	TTT-TCA-GGT-TTA-CCT-TGA-GC
fumC		
CJfumC1(+)	A, S	TCG-TGC-CAC-TGA-AAT-CAT-GG
CJfumC2(−)	A, S	ACC-TAT-GTG-TGG-ATT-TAG-AGC
nuoH		
Cjnadh1(+)	A, S	GCA-GCT-ATT-CCT-ATG-CTA-CC
CJnadh2(−)	A, S	TTG-ATC-TGG-ACG-CAA-TTG-CG
yphC		
CJyphc1(+)	A, S	TAT-CAG-AGT-GGG-TAT-TGT-AGG
CJyphc2(−)	A, S	AAT-CAC-TAA-AGG-CAC-ACC-TTC

[a] A, amplification; S, sequencing.

genes, are widely separated on the chromosome, and are not located in the vicinity of putative virulence genes or outer membrane protein genes. PCR amplification and direct sequencing of PCR products were performed as described previously (30). Briefly, total DNA was purified with the QiaAmp tissue kit (Qiagen). PCR products were generated with the primers listed in Table 3 and sequenced from both strands on an ABI 377 automated sequencer.

Phylogenetic analysis. Sequences were aligned by using SEQLAB and PILEUP from the Genetics Computer Group (Madison, Wis.) Wisconsin Package, version 9.1. All sequences for one gene fragment were reduced to a common length and exported to MSF (multisequence file) format. Where necessary, sequences were converted to an MEG (MEGA) format with the program PSFIND (kindly provided by Mark Achtman). K_A and K_S values with Jukes-Cantor corrections were calculated with DNASP 3.0 (26). The homoplasy test (21) was performed with HOMOPLASY (30). The sequence alignments were converted to NEXUS files by using SFE 1.0.3 (K. Jolley, http://mlst.zoo.ox.ac.uk/links/SFE103.zip), and split decomposition was analyzed with SPLITSTREE 3.1 (17). Allele numbers were assigned with SEQUENCE OUTPUT (B. G. Spratt, http://mlst.zoo.ox.ac.uk/links/SeqOutput.sit). The standardized index of association (sI_A)(15) was calculated with LIAN 3.0 (http://seneca.ice.mpg.de/lian) (14). The UPGMA (unweighted pair group mean average) tree shown in Fig. 1 was drawn with START (K. Jolley, http://mlst.zoo.ox.ac.uk/links/START.zip).

Nucleotide sequence accession number. The sequences of all alleles have been deposited in the EMBL/GenBank databases, and the accession numbers are listed in Table 2.

RESULTS

Sequence diversity in *Campylobacter jejuni*. Seven housekeeping gene fragments were selected for this analysis on the basis of (i) being distributed widely on the chromosome of *C. jejuni* NCTC 11168 and (ii) the assumption that the vicinity of

TABLE 2. Gene fragments that were sequenced

Gene	Gene numbering[a]	Position[a]	Length (bp)[b]	Gene product	Accession no.
asd	Cj1023c	954448	564	Aspartate-semialdehyde dehydrogenase	AJ292175–AJ292188
atpA	Cj0105	111788	660	ATP synthase, F1α	AJ292166–AJ292174
ddlA	Cj0798c	748473	480	D-Alanine-D-alanine-ligase	AJ290352–AJ290363
eftS (*tsf*)	Cj1181c	1108052	444	Elongation factor TS	AJ290337–AJ290351
fumC	Cj1364c	1296454	645	Fumarate hydratase	AJ290322–AJ290336
nuoH	Cj1572c	1502750	423	NADH dehydrogenase H	AJ290377–AJ290387
yphC	Cj0386	352112	570	GTP-binding protein	AJ290364–AJ290476

[a] The gene numbering and the indication of the gene position refer to the genome sequence of *C. jejuni* NCTC 11168 published by the Sanger Centre (http://www.sanger.ac.uk/Projects/C_jejuni/).
[b] Length of sequenced fragment.

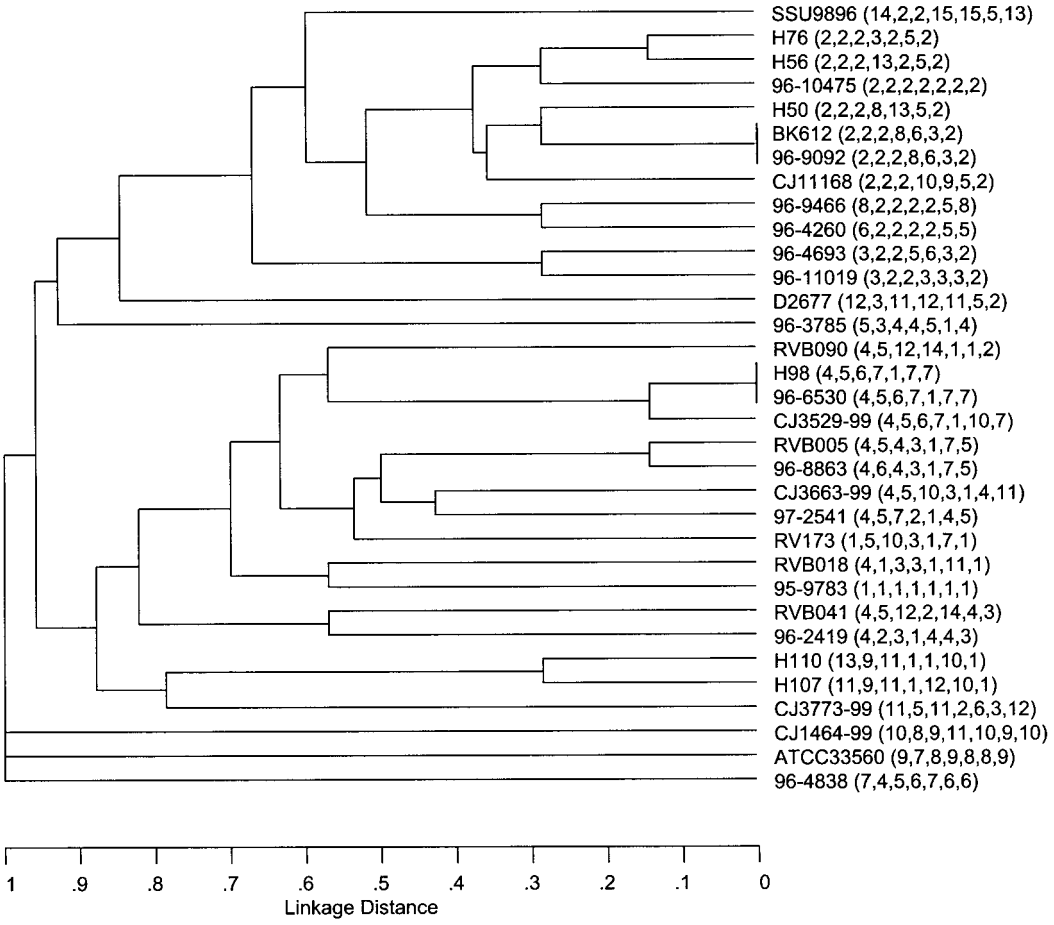

FIG. 1. UPGMA dendrogram showing the genetic relatedness of the 33 *C. jejuni* strains examined in this study. The dendrogram was constructed from a matrix of the pairwise distances between the allelic profiles of 33 *C. jejuni* strains. The numbers in parentheses behind the strain name represent the allele profile for that strain. Allele numbers are listed in the following order: *asd*, *atpA*, *ddlA*, *eftS*, *fumC*, *nuoH*, and *yphC*.

the genes did not contain virulence genes or other genes (such as outer membrane protein genes) that could be predicted to be under strong selection (Table 2). Primers were designed to amplify fragments internal to these housekeeping genes, and the PCR products were sequenced from both strands. The strains used were isolated from patients with enteritis in two regions of Germany (Freiburg, 13 strains; Würzburg, 5 strains), Hungary (6 strains), and Thailand (5 strains) and included 2 strains from the United States and the *C. jejuni* subsp. *jejuni* type strain, ATCC 33560. In addition, the sequences for *C. jejuni* strain NCTC 11168, whose genome sequence was recently published, were added to all data sets. The 13 strains from the Freiburg region were selected from a larger collection of strains from 102 patients to be as diverse as possible, based on characterizations by pulsed-field gel electrophoresis analysis. The strains from the Würzburg region were random clinical isolates from patients with enteritis and one patient with *C. jejuni* bacteremia (BK612). The strains from Thailand were isolated in Children's Hospital, Bangkok, and were epidemiologically unrelated according to isolation date and serotype. The strains from Hungary were collected in Budapest and from two counties of Hungary (Pest and Veszprem) and were isolated from sporadic enteritis cases.

For all loci, all 33 sequences aligned without gaps or insertions. For all genes, multiple strains carried identical alleles. The number of different sequences (alleles) varied between 9 for *atpA* and 15 for *eftS* and *fumC* (Table 4). However, only two pairs of strains had identical sequences for all seven fragments, and all other strains could be distinguished from each other because they had unique combinations of alleles. The allelic profiles and a UPGMA tree based on these allelic profiles are shown in Fig. 1. There were no significant clusters that correlated with either geographic origin of the strains or serotype.

As expected for housekeeping genes, the frequency of nonsynonymous mutations leading to changes of the amino acid sequence was quite low (K_A below 1% for all fragments). The frequency of synonymous mutations varied between 0.86% for *atpA* and 8.94% for *fumC*, and the K_S/K_A ratio varied between 5.84 (*nuoH*) and 72 (*atpA*).

Sequence diversity in *atpA* of *C. jejuni* and *Helicobacter pylori*. One of the fragments sequenced overlapped with a fragment we previously sequenced for 20 strains of *H. pylori* (2). A fragment of 591 bp was available for all 33 *C. jejuni* strains and 20 strains of *H. pylori*. While all 20 *H. pylori* sequences were unique, there were only 9 different alleles among the 33 *C. jejuni* strains. The frequency of both synonymous and nonsyn-

TABLE 4. Allelic diversity and recombination in housekeeping genes of *C. jejuni*

Gene	No. of alleles	No. of polymorphic sites	Mean % (range):		K_S/K_A ratio	*H* ratio[a]
			K_S	K_A		
asd	14	25	3.8 ± 2.6 (0–10.3)	0.4 ± 0.3 (0–0.9)	12.1	0.40
atpA	9	11	0.86 ± 0.73 (0–3.8)	0.01 ± 0.05 (0–0.2)	71.5	
ddlA	12	39	8.9 ± 5.3 (0–20.5)	0.8 ± 0.6 (0–2.0)	11.6	0.42
eftS (*tsf*)	15	28	7.1 ± 5.2 (0–19.5)	0.2 ± 0.2 (0–1.0)	32.3	0.36
fumC	15	39	6.4 ± 5.4 (0–19.3)	0.14 ± 0.14 (0–0.4)	45.7	0.48
nuoH	11	23	3.18 ± 2.53 (0–9.35)	0.6 ± 0.4 (0–1.3)	5.8	
yphC	13	23	1.9 ± 2.3 (0–10.0)	0.1 ± 0.2 (0–0.7)	16.0	0.47

[a] Geometric mean of 10 independent runs of the homoplasy test. The standard error was <10% for all values. *H* ratios for *atpA* and *nuoH* could not be determined because there were not enough informative sites.

onymous mutations in this fragment was much lower in *C. jejuni* than in *H. pylori*. The K_S and K_A for *H. pylori* were 12.3% and 0.26%, respectively, compared with 0.84% and 0.014% for *C. jejuni*. A neighbor-joining tree (Jukes-Cantor distances) of these sequences separated all *H. pylori* sequences from all *C. jejuni* sequences in two widely separated branches. Within these two branches, there were no significant further branches, as assessed by a bootstrap test (not shown).

Evidence for recombination in *C. jejuni*. Bacterial pathogens can differ widely in their population structure. While some (e.g., *Borrelia burgdorferi*) have a clonal population structure, in which all sequence diversity has arisen by sequential accumulation of point mutations, in many other species (e.g., *Neisseria meningitidis* and *H. pylori*), sequence diversity is greatly increased by intraspecies recombination (10, 22, 30). The sets of sequences were tested with the homoplasy test (21), which analyzes the apparent homoplasies among informative, synonymous polymorphic sites. The frequency of apparent homoplasies, as measured by the homoplasy ratio, *H*, is an indicator of the frequency of recombination. The homoplasy ratio can vary between 0, indicating completely clonal descent by the accumulation of mutations, and 1, indicating free recombination where all sequence polymorphisms are found repeatedly in independent sequences that are in different branches of a maximal parsimony tree (apparent homoplasies). The homoplasy test requires a sufficient number of informative sites to yield interpretable results. This was only the case for five out of the seven fragments. In the *atpA* and *nuoH* data sets, there was not sufficient sequence diversity to perform the homoplasy test. For the remaining five loci, the homoplasy test gave *H* ratios between 0.36 and 0.48, indicating frequent intraspecific recombination (Table 4).

Different genes show network-like and tree-like phylogeny in *C. jejuni*. In order to further analyze the population structure of *C. jejuni*, we used split decomposition analysis. Split decomposition was originally described by Bandelt and Dress in 1992 (5) and has since been used to analyze the population structures of both bacteria and viruses (7, 16, 28). Because this method does not make the a priori assumption that the sequences have a tree-like structure, conflicting phylogenetic signals in the data, such as evidence of recombination, can be visualized and will lead to the generation of an interconnected network rather than a tree. Figure 2 shows the split graphs for all alleles of the seven fragments analyzed. The structure of the split graphs varied substantially between the different loci. Four of the seven loci contained a more (*asd* and *yphC)* or less (*eftS* and *nuoH*) complex network. The split graphs of the other three loci showed no evidence of network-like evolution. The fit parameter was 100 for *asd*, *atpA*, *nuoH*, and *yphC*, indicating that all phylogenetic information in the sequences could be visualized in the graphs. The fit parameters for the other three loci (*ddlA*, *eftS*, and *fumC*) were lower (59.7, 41.4, and 47.6, respectively), indicating that not all information could be integrated into the graph.

These analyses are consistent with the results of the homoplasy tests in that recombination has occurred within most the loci analyzed and contributed to the generation of sequence diversity. The differences in structure between the graphs obtained for the seven loci can also best be explained by recombination, because recombination can lead to the assembly of genes with different evolutionary histories within one strain. A split graph analysis was also performed based on a distance matrix of pairwise distances of the allelic profiles of all strains (Fig. 3). This graph resembled a star with rays of different lengths, consistent with a recombinational population structure (star phylogeny).

Statistical testing for linkage disequilibrium and sI_A. The index of association (I_A) (22) has been widely used to analyze the degree of linkage in multilocus enzyme electrophoresis data sets. I_A, as originally described by Maynard Smith, is dependent upon the number of loci analyzed (20), which is taken into account in the calculation of the standardized I_A (sI_A) (14, 15). sI_A is expected to be 0 when alleles are in linkage equilibrium because of free recombination. sI_A was 0.256, consistent with a limited amount of recombination that did not completely destroy the linkage between alleles. The difference from linkage equilibrium was statistically highly significant (test by Monte Carlo simulation, 100,000 repetitions, $P < 10^{-5}$).

DISCUSSION

The results presented here show that *C. jejuni* has a relatively small pool of unique alleles with few polymorphic nucleotides. We initially only studied strains from Germany and included a small set of previously characterized strains from the United States and the United Kingdom. The subsequent addition of 11 strains from Hungary and Thailand only increased the number of unique alleles by one or two per locus, despite the fact that the first sample was from a relatively small

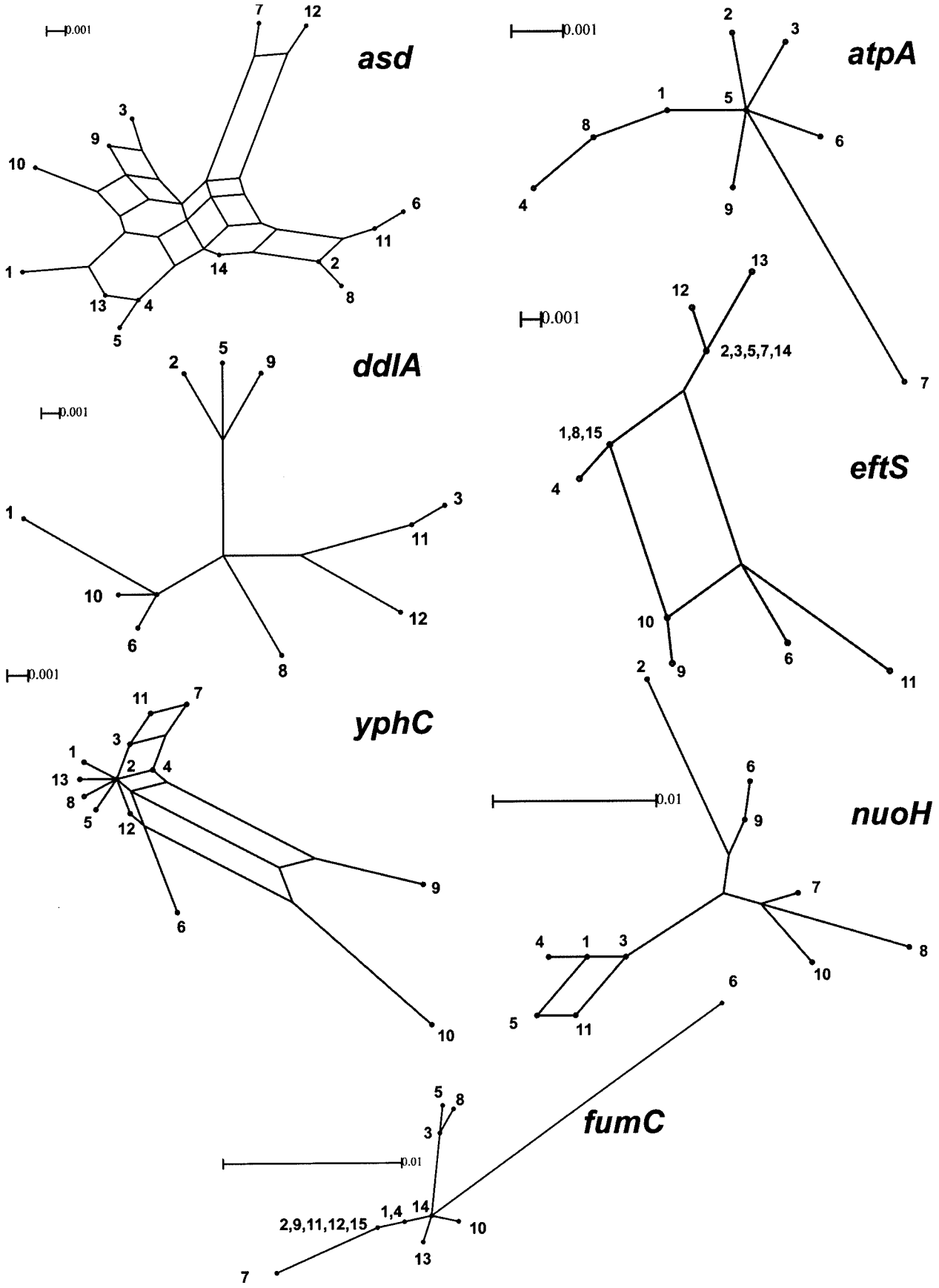

FIG. 2. Split decomposition analysis of 9 to15 unique alleles obtained from 33 C. jejuni strains for seven loci. All branch lengths are drawn to scale. The observation that in the *asd*, *yphC*, *nuoH*, and *eftS* graphs several alleles in the sample are connected to each other by multiple pathways, forming an interconnected network, is suggestive of recombination. The numbering refers to allele numbers (Fig. 1).

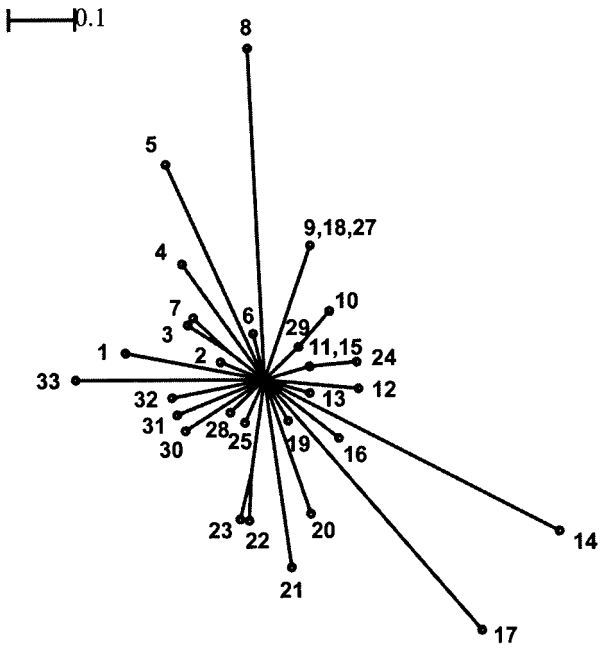

FIG. 3. Split decomposition analysis based on a matrix of pairwise distances between the allelic profiles of the 33 *C. jejuni* strains examined in this study. The numbering refers to strain numbers (Table 1).

geographic area. However, these sequences occur in many different combinations, so that we observed 31 unique sequence types among the 33 strains. The two pairs of strains that had identical allele profiles for the seven genes analyzed here could easily be differentiated by either RAPD PCR or sequencing of a highly polymorphic fragment that included the intergenic region between *adk* and *ppa* (data not shown).

Both the homoplasy test and the split decomposition analyses give strong evidence that intraspecies recombination occurs frequently in *C. jejuni* and plays a major role in generating genetic heterogeneity between strains. The mean H ratio observed for the five loci analyzed (0.42 ± 0.05) was higher than values previously reported for *Neisseria meningitidis* (11 genes, mean $H = 0.34$) (30), *Escherichia coli* (four genes, mean $H = 0.26$) (30), *S. pneumoniae* (2 genes, mean $H = 0.3$) (29), and *Borrelia burgdorferi* (1 gene, $H = 0.06$) (21). The only organism in which higher homoplasy ratios were reported is *H. pylori* (seven housekeeping genes, mean $H = 0.65 \pm 0.11$) (2). The observation of recombination in *C. jejuni* is in agreement with previous analyses of the two tandem flagellin genes *flaA* and *flaB,* in which both intragenomic recombination and interstrain recombination have been shown to occur (13, 33).

Because *C. jejuni* and *H. pylori* are phylogenetically closely related gastrointestinal pathogens, these two warrant a more detailed comparison. Both species share features that would be expected have an influence on population structure. Specifically, both have comparatively small, AT-rich genomes, are naturally competent for DNA uptake, and contain abundant hypermutatable simple nucleotide repeats that permit switching of genes on or off (25, 27, 31). However, the basic features of the population structure of *C. jejuni* described here differ markedly from those of *H. pylori*. In *H. pylori*, it is most unusual to find two unrelated strains with the same nucleotide sequence in any given gene (18, 30). In a multilocus sequencing study of 20 *H. pylori* strains that included seven housekeeping genes, the number of unique alleles per locus was between 18 and 20 (2). Available data suggest that, in *H. pylori,* frequent recombination in the absence of effective purification mechanisms (such as sequential bottlenecks or founder effects that would eliminate recombinant genotypes from the population) is responsible for creating an almost unlimited number of unique alleles. The data reported here for *C. jejuni* show that the frequency of recombination (or, more precisely, the probability that recombinant genotypes remain in the population) is lower than in *H. pylori*, where individual polymorphic sites were close to linkage equilibrium, effectively creating random combinations of available polymorphic nucleotides (30). Whether this is due to a lower frequency of DNA transfer events or to ecological differences that imply more effective purification mechanisms is not known. However, recombination in *C. jejuni* occurs frequently enough to create many different combinations of alleles (unique sequence types), as reflected by the low I_A and the large number of unique sequence types. A practical implication of this finding is that an MLST approach therefore bears great promise to be an efficient typing method for large-scale use in *C. jejuni* molecular epidemiology, despite the overall very low level of sequence diversity. There was no obvious association between strains with similar allelic profiles and either geographic origin or serotype. However, serotyping data were not available for all strains, and the number of strains for any serotype was too small to allow definite conclusions. MLST analyses of larger groups of strains with different serotypes will be necessary to define the relationship between serotypes and MLST allelic profiles. In fact, a very recently published MLST study of 154 strains of *C. jejuni* coming almost exclusively from the United Kingdom has provided evidence for an association of some serotypes with certain clonal lineages (6).

Why was there so little sequence diversity in the *C. jejuni* isolates studied, which almost exclusively came from infected humans? One possibility is that *C. jejuni* is a young species and therefore has not yet had enough time to accumulate sequence diversity. Total lack of sequence diversity in the etiological agent of plague, *Y. pestis*, has recently been explained by a very recent emergence of this "species" as a new clone of *Yersinia pseudotuberculosis* (3). Another (and maybe the most likely) explanation for the relative paucity of sequence variation is that the *C. jejuni* population has recently undergone rapid expansion, possibly driven by the changes in food animal husbandry and slaughtering practices in the last one or two centuries. Such an expansion of clones particularly fit to survive under the conditions of industrialized animal husbandry could explain the limited number of alleles and the low frequency of synonymous nucleotide polymorphisms.

Frequent recombination between strains is likely to greatly facilitate the spread of favorable traits such as antibiotic resistance genes in the population and may have contributed to the rapid global spread of resistance against macrolides and gyrase inhibitors in this important pathogen.

ACKNOWLEDGMENTS

We gratefully acknowledge the *Campylobacter* genome sequencing project at the Sanger Centre. This project profited enormously from their liberal data release policy. We thank Rick Meinermann for providing DNAs from several *C. jejuni* strains and Lorrin Pang (Armed Forces Research Institute for Medical Sciences, Bangkok) for providing the strains from Thailand. We are grateful to Mark Achtman and Christine Josenhaus for critical reading of the manuscript, Bernhard Haubold for helpful advice about the use of LIAN, Matthias Frosch, Helge Karch, and Bernhardt Steinbrückner for helpful support, and Barbara Beuerle for excellent technical assistance.

This work was supported by grants Su 133/2–3 (Gerhard Hess award) and Su 133/3–1 from the Deutsche Forschungsgemeinschaft to S.S. and by a Deutscher Akademikeraustauschdienst (DAAD) fellowship to A.S.

REFERENCES

1. **Achtman, M.** 2001. Population structure of *Helicobacter pylori* and other pathogenic bacterial species, p. 311–321. *In* M. Achtman and S. Suerbaum (ed.), *Helicobacter pylori*: molecular and cellular biology. Horizon Scientific Press, Wymondham, United Kingdom.
2. **Achtman, M., T. Azuma, D. E. Berg, Y. Ito, G. Morelli, Z. J. Pan, S. Suerbaum, S. A. Thompson, A. van der Ende, and L. J. van Doorn.** 1999. Recombination and clonal groupings within *Helicobacter pylori* from different geographic regions. Mol. Microbiol. **32:**459–470.
3. **Achtman, M., K. Zurth, G. Morelli, G. Torrea, and E. Carniel.** 1999. *Yersinia pestis*, the cause of pandemic plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. Proc. Natl. Acad. Sci. USA **96:**14043–14048.
4. **Alm, R. A., P. Guerry, and T. J. Trust.** 1993. Distribution and polymorphism of the flagellin genes from isolates of *Campylobacter coli* and *Campylobacter jejuni*. J. Bacteriol. **175:**3051–3057.
5. **Bandelt, H. J., and A. W. Dress.** 1992. Split decomposition: a new and useful approach to phylogenetic analysis of distance data. Mol. Phylogenet. Evol. **1:**242–252.
6. **Dingle, K. E., F. M. Colles, D. R. Wareing, R. Ure, A. J. Fox, F. E. Bolton, H. J. Bootsma, R. J. Willems, R. Urwin, and M. C. Maiden.** 2001. Multilocus sequence typing system for *Campylobacter jejuni*. J. Clin. Microbiol. **39:**14–23.
7. **Dopazo, J., A. Dress, and A. von Haeseler.** 1993. Split decomposition: a technique to analyze viral evolution. Proc. Natl. Acad. Sci. USA **90:**10320–10324.
8. **Duim, B., T. M. Wassenaar, A. Rigter, and J. Wagenaar.** 1999. High-resolution genotyping of *Campylobacter* strains isolated from poultry and humans with amplified fragment length polymorphism fingerprinting. Appl. Environ. Microbiol. **65:**2369–2375.
9. **Enright, M. C., and B. G. Spratt.** 1998. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. Microbiology **144:**3049–3060.
10. **Feil, E. J., M. C. Maiden, M. Achtman, and B. G. Spratt.** 1999. The relative contributions of recombination and mutation to the divergence of clones of *Neisseria meningitidis*. Mol. Biol. Evol. **16:**1496–1502.
11. **Fujimoto, S., B. M. Allos, N. Misawa, C. M. Patton, and M. J. Blaser.** 1997. Restriction fragment length polymorphism analysis and random amplified polymorphic DNA analysis of *Campylobacter jejuni* strains isolated from patients with Guillain-Barre syndrome. J. Infect. Dis. **176:**1105–1108.
12. **Gibson, J. R., C. Fitzgerald, and R. J. Owen.** 1995. Comparison of PFGE, ribotyping and phage-typing in the epidemiological analysis of *Campylobacter jejuni* serotype HS2 infections. Epidemiol. Infect. **115:**215–225.
13. **Harrington, C. S., F. M. Thomson-Carter, and P. E. Carter.** 1997. Evidence for recombination in the flagellin locus of *Campylobacter jejuni*: implications for the flagellin gene typing scheme. J. Clin. Microbiol. **35:**2386–2392.
14. **Haubold, B., and R. R. Hudson.** 2000. LIAN 3.0: detecting linkage disequilibrium in multilocus data. Bioinformatics **16:**847–849.
15. **Haubold, B., M. Travisano, P. B. Rainey, and R. R. Hudson.** 1998. Detecting linkage disequilibrium in bacterial populations. Genetics **150:**1341–1348.
16. **Holmes, E. C., M. Worobey, and A. Rambaut.** 1999. Phylogenetic evidence for recombination in dengue virus. Mol. Biol. Evol. **16:**405–409.
17. **Huson, D. H.** 1998. SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics **14:**68–73.
18. **Kansau, I., J. Raymond, E. Bingen, P. Courcoux, N. Kalach, M. Bergeret, N. Braimi, C. Dupont, and A. Labigne.** 1996. Genotyping of *Helicobacter pylori* isolates by sequencing of PCR products and comparison with the RAPD technique. Res. Microbiol. **147:**661–669.
19. **Maiden, M. C., J. A. Bygraves, E. Feil, G. Morelli, J. E. Russell, R. Urwin, Q. Zhang, J. Zhou, K. Zurth, D. A. Caugant, I. M. Feavers, M. Achtman, and B. G. Spratt.** 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. USA **95:**3140–3145.
20. **Maynard Smith, J.** 1999. The detection and measurement of recombination from sequence data. Genetics **153:**1021–1027.
21. **Maynard Smith, J., and N. H. Smith.** 1998. Detecting recombination from gene trees. Mol. Biol. Evol. **15:**590–599.
22. **Maynard Smith, J., N. H. Smith, M. O'Rourke, and B. G. Spratt.** 1993. How clonal are bacteria? Proc. Natl. Acad. Sci. USA **90:**4384–4388.
23. **Mazurier, S., A. van de Giessen, K. Heuvelman, and K. Wernars.** 1992. RAPD analysis of *Campylobacter* isolates: DNA fingerprinting without the need to purify DNA. Lett. Appl. Microbiol. **14:**260–262.
24. **Meinersmann, R. J., L. O. Helsel, P. I. Fields, and K. L. Hiett.** 1997. Discrimination of *Campylobacter jejuni* isolates by *fla* gene sequencing. J. Clin. Microbiol. **35:**2810–2814.
25. **Parkhill, J., B. W. Wren, K. Mungall, J. M. Ketley, C. Churcher, D. Basham, T. Chillingworth, R. M. Davies, T. Feltwell, S. Holroyd, K. Jagels, A. V. Karlyshev, S. Moule, M. J. Pallen, C. W. Penn, M. A. Quail, M. A. Rajandream, K. M. Rutherford, A. H. van Vliet, S. Whitehead, and B. G. Barrell.** 2000. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. Nature **403:**665–668.
26. **Rozas, J., and R. Rozas.** 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics **15:**174–175.
27. **Saunders, N. J., J. F. Peden, D. W. Hood, and E. R. Moxon.** 1998. Simple sequence repeats in the *Helicobacter pylori* genome. Mol. Microbiol. **27:**1091–1098.
28. **Smith, N. H., E. C. Holmes, G. M. Donovan, G. A. Carpenter, and B. G. Spratt.** 1999. Networks and groups within the genus *Neisseria*: analysis of *argF*, *recA*, *rho*, and 16S rRNA sequences from human *Neisseria* species. Mol. Biol. Evol. **16:**773–783.
29. **Suerbaum, S.** 2000. Genetic variability within *Helicobacter pylori*. Int. J. Med. Microbiol. **290:**175–181.
30. **Suerbaum, S., J. Maynard Smith, K. Bapumia, G. Morelli, N. H. Smith, E. Kunstmann, I. Dyrek, and M. Achtman.** 1998. Free recombination within *Helicobacter pylori*. Proc. Natl. Acad. Sci. USA **95:**12619–12624.
31. **Tomb, J.-F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann, K. A. Ketchum, H. P. Klenk, S. Gill, B. A. Dougherty, K. Nelson, J. Quackenbush, L. Zhou, E. F. Kirkness, S. Peterson, B. Loftus, D. Richardson, R. Dodson, H. G. Khalak, A. Glodek, K. McKenney, L. M. Fitzgerald, N. Lee, M. D. Adams, E. K. Hickey, D. E. Berg, J. D. Gocayne, T. R. Utterback, J. D. Peterson, J. M. Kelley, M. D. Cotton, J. M. Weidman, C. Fujii, C. Bowman, L. Watthey, E. Wallin, W. S. Hayes, M. Borodovsky, P. D. Karp, H. O. Smith, C. M. Fraser, and J. C. Venter.** 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature **388:**539–547.
32. **Wassenaar, T. M., and M. J. Blaser.** 1999. Pathophysiology of *Campylobacter jejuni* infections of humans. Microbes Infect **1:**1023–1033.
33. **Wassenaar, T. M., B. N. Fry, and B. A. van der Zeijst.** 1995. Variation of the flagellin gene locus of *Campylobacter jejuni* by recombination and horizontal gene transfer. Microbiology **141:**95–101.
34. **Wassenaar, T. M., and D. G. Newell.** 2000. Genotyping of *Campylobacter* spp. Appl. Environ. Microbiol. **66:**1–9.
35. **Yan, W., N. Chang, and D. E. Taylor.** 1991. Pulsed-field gel electrophoresis of *Campylobacter jejuni* and *Campylobacter coli* genomic DNA and its epidemiologic application. J. Infect. Dis. **163:**1068–1072.