The Genome of *Salmonella enterica* Serovar Gallinarum: Distinct Insertions/Deletions and Rare Rearrangements

Kai-Yu Wu,¹ Gui-Rong Liu,^{1,2} Wei-Qiao Liu,³ Austin Q. Wang,¹ Sen Zhan,¹
Kenneth E. Sanderson,³ Randal N. Johnston,⁴ and Shu-Lin Liu^{1,2*}

Departments of Microbiology and Infectious Diseases,¹ Biological Sciences,³ and Biochemistry and Molecular Biology,⁴ University of Calgary, Calgary, Alberta, Canada, and Department of Microbiology, Peking University Health Science Center, Beijing, China²

Received 4 October 2004/Accepted 19 April 2005

Salmonella enterica serovar Gallinarum is a fowl-adapted pathogen, causing typhoid fever in chickens. It has the same antigenic formula (1,9,12:—:—) as *S. enterica* serovar Pullorum, which is also adapted to fowl but causes pullorum disease (diarrhea). The close relatedness but distinct pathogeneses make this pair of fowl pathogens good models for studies of bacterial genomic evolution and the way these organisms acquired pathogenicity. To locate and characterize the genomic differences between serovar Gallinarum and other salmonellae, we constructed a physical map of serovar Gallinarum strain SARB21 by using I-CeuI, XbaI, and AvrII with pulsed-field gel electrophoresis techniques. In the 4,740-kb genome, we located two insertions and six deletions relative to the genome of *S. enterica* serovar Typhimurium LT2, which we used as a reference *Salmonella* genome. Four of the genomic regions with reduced lengths corresponded to the four prophages in the genome of serovar Typhimurium LT2, and the others contained several smaller deletions relative to serovar Typhimurium LT2, including regions containing *srfJ*, *std*, and *stj* and gene clusters encoding a type I restriction system in serovar Typhimurium LT2. The map also revealed some rare rearrangements, including two inversions and several translocations. Further characterization of these insertions, deletions, and rearrangements will provide new insights into the molecular basis for the specific host-pathogen interactions and mechanisms of genomic evolution to create a new pathogen.

Salmonella species are important bacterial pathogens. To date, more than 2,500 *Salmonella* serovars have been identified (39). Genetic evidence shows that all *Salmonella* serovars are closely related (3, 16). For example, comparisons of house-keeping genes indicate 96 to 99% sequence identity among the *Salmonella* serovars (5). According to microarray data, 2,244 potential genes are present in all 24 examined strains of 12 *S. enterica* serovars and 2 strains of *S. bongori* (2); similar observations have also been made by Porwollik and his colleagues (40). Such genes are termed core genes and are assumed to play key roles in many typical cellular functions, e.g., metabolism, biosynthesis, and DNA replication. On the other hand, although the *Salmonella* serovars are closely related by phylogeny, they differ significantly in biology, including factors such as host range and disease spectrum. For instance, serovar Typhimurium infects a broad range of hosts, causing enteritis in humans and typhoid fever in mice, whereas serovar Typhi infects only humans and causes only typhoid fever. It is thus of great interest to ask what genetic differences in the *Salmonella* genomes have caused such phenotypic diversity among these close relatives.

Genomic comparisons between different *Salmonella* serovars have revealed that there are numerous serovar-specific insertions, deletions, inversions, translocations, and pseudogenes in the genomes (4, 20, 22–26, 30, 34, 35, 38). The insertions usually have G+C contents that differ from the average

for the whole genome, an indication of exogenous origin. Some insertions have been designated *Salmonella* pathogenicity islands (SPIs) because they contain virulence genes contributing to bacterial pathogenicity. SPI-7, one of the recently identified SPIs, for instance, encodes Vi antigen in serovar Typhi (36, 38). Genomic sequence analysis has revealed many pseudogenes in the genome of serovar Typhi, which may account at least in part for its host specificity.

Serovars Gallinarum and Pullorum are fowl-adapted pathogens and the only known nonflagellate (and thus nonmotile) *Salmonella* lineages. They have the same antigenic formula (1,9,12:—:—) and are indeed very closely related. However, they cause very different diseases, with serovar Gallinarum causing fowl typhoid fever and serovar Pullorum causing pullorum disease. Fowl typhoid fever is a chronic disease of adult chickens, but pullorum disease is an acute diarrheal illness in chicks. Fowl typhoid fever is usually spread by the ingestion of contaminated food or water, whereas pullorum disease can be transferred vertically through eggs in carrier hens (17, 43). The close relatedness, common host specificity, and different pathogeneses of serovars Gallinarum and Pullorum make them good models in comparative genomics for exploration of the evolutionary events converting nonpathogenic bacteria into different pathogens. It is reasonable to suppose that some genomic features shared only by the two *Salmonella* lineages may be responsible for their common phenotypes, such as host range, and other genomic features unique to each of them may account for their distinct pathogeneses. In this study, we aim at finding such respective common and distinct genomic features of serovars Gallinarum and Pullorum.

* Corresponding author. Mailing address: Department of Microbiology and Infectious Diseases, University of Calgary, 3330 Hospital Dr. NW, Calgary, Alberta T2N 4N1, Canada. Phone: (403) 220-3799. Fax: (403) 270-0834. E-mail: slliu@ucalgary.ca.

TABLE 1. Primers used for determining the orientations of the I-CeuI fragments in serovar Gallinarum SARB21

Primer	Gene	STY no. ^a	Position (nt) ^b of gene in serovar Typhi CT18	Primer location (nt)	Sequence
1	23S rRNA		4259388 to 4262393	4262300 to 4262319	GGAACGTTGAAGACGACGAC
2	16S rRNA		4257492 to 4259033	4257578 to 4257559	GCAAGCTGCTTCCTGTTACC
3	STY 2848	2848	2711380 to 2711835	2711789 to 2711809	TCAATGCTGGAAAAGTCTTGC
4	<i>clpB</i>	2849	2717996 to 2720569	2718060 to 2718041	AACTGGTTCTCGGCAAAGTG
5	<i>murB</i>	3742	3593509 to 3594537	3594440 to 3594459	CAGGCGCTCAGTAGTTGTTG
6	<i>murI</i>	3743	3600451 to 3601233	3600501 to 3600482	CCCGTTTACAGCGTTATGG
7	<i>yafB</i>	276	292894 to 293697	292915 to 292896	AACCGAATGCAGGGATAGTC
8	<i>yaeD</i>	275	286542 to 287108	286982 to 287001	AAATGTCGGGACAAAAGTGC
9	<i>yifA</i>	3708	3551580 to 3552416	3552341 to 3552359	CATTTGGCTGCAAAACAGC
10	<i>purH</i>	3709	3558479 to 3560068	3558553 to 3558534	CTGGGCGAATTCGATGATAC
11	<i>yjaB</i>	4399	4262945 to 4263382	4262985 to 4262966	TTGGGAAACCGTATCCATTG
12	<i>yrdA</i>	4398	4256466 to 4257020	4256923 to 4256942	AGACGGACATCGCCAATAAC
13	<i>mobB</i>	3887	3743139 to 3743654	3743624 to 3743643	GCGGTGGTTAGAGAAAGCAC
14	<i>yieP</i>	3888	3749553 to 3750257	3749676 to 3749657	CTATTTCGCGGGGAAGAATC
15	<i>yhdV</i>	3572	3416892 to 3417113	3417087 to 3417106	ATGGGTATCATTGCGCTTTC
16	<i>hemG</i>	3573	3423822 to 3424367	3424274 to 3424256	GCCGATGTCGTTAATCTGC

^a STY no. represents the genes in *S. enterica* serovar Typhi strain CT18 (38).

^b nt, nucleotides.

To this end, we mapped the genome of serovar Gallinarum SARB21 and compared it to the previously mapped genomes of serovars Typhimurium (23) and Pullorum (19). This map and its comparisons with those of the genomes of the other *Salmonella* serovars revealed several insertions, deletions, and rearrangements in serovar Gallinarum SARB21. The located deletions were further identified and characterized by using bioinformatics tools.

MATERIALS AND METHODS

Bacterial strains and cultivation conditions. Serovar Gallinarum SARB21, originally isolated as CDC4801/72, was kindly provided by R. K. Selander and was stocked at the *Salmonella* Genetic Stock Center as SGSC2478 (<http://www.ucalgary.ca/~kesander>). The bacteria were grown in Luria-Bertani (LB) broth or on LB plates at 37°C. Tetracycline was used at 20 µg/ml to select for Tn10 insertion mutants.

Enzymes and chemicals. I-CeuI, XbaI, and AvrII were purchased from New England Biolabs, proteinase K was from Boehringer Mannheim, and [³²P]dCTP was from Amersham Pharmacia Biotech. Most other chemicals were from Sigma Chemical Co.

Generation of Tn10 insertion mutants through bacteriophage P22-mediated transduction. A large number of Tn10 insertions into known genes have been mapped for the genome of serovar Typhimurium LT2 (23). Using bacteriophage P22, we transferred Tn10 insertions in 60 representative genes from the serovar Typhimurium LT2 background into the genome of serovar Gallinarum SARB21 by the procedures described previously (22).

PFGE methods and genomic mapping. Pulsed-field gel electrophoresis (PFGE) techniques, including preparation of genomic DNA and separation of the cleaved DNA fragments by PFGE, were as described previously (28, 29).

Long-range PCR to determine orientations of I-CeuI fragments. Primers used in long-range PCR to determine the orientations of I-CeuI fragments are listed in Table 1, and the procedure was as described previously (18); a similar procedure has been used by Helm and Maloy (8).

Bioinformatics tools and analyses. *Salmonella* sequences were downloaded from the Genome Sequencing Center at Washington University (<http://genome.wustl.edu>). Genes that were not present in serovar Gallinarum as suggested by the PFGE data were identified using BLAST searching against shotgun sequences of another strain of serovar Gallinarum, 287/91, provided by the Sanger Center (<http://www.sanger.ac.uk/cgi-bin/blast/submitblast/salmonella>). Generally, the sequence identities of orthologous genes identified in BLAST searches were >96% among the *Salmonella* lineages in our comparisons. Genes were regarded as missing when sequence identities were less than 70%.

RESULTS

I-CeuI cleavage map of serovar Gallinarum SARB21.

I-CeuI, an endonuclease that cleaves DNA in a 26-bp recognition site in bacterial *rhl* genes (6, 32, 33), cleaves *Escherichia coli* or *Salmonella* genomes into seven fragments because these bacteria all have seven *rnm* operons (21). We arbitrarily designated these fragments in serovar Typhimurium LT2 in the order ABCDEFG, starting from the fragment that contains the terminus of replication and going clockwise (23). In most *Salmonella* serovars, the seven I-CeuI fragments are organized in the same order as those in serovar Typhimurium LT2 (21, 27, 31). As shown in Fig. 1, serovar Gallinarum SARB21 has a rearranged genome, with the seven I-CeuI fragments being organized differently than those in serovar Typhimurium LT2, a situation that has been seen in a number of other host-adapted *Salmonella* serovars (19, 25, 30). Although the partial digestion products, such as E+F, E+D, F+C, and etc. (Fig. 1A), clearly demonstrate the order of the seven I-CeuI fragments to be ABDEFCEG in serovar Gallinarum SARB21 (Fig. 1B), the orientation of the fragments, especially A and C, could not be determined this way. By long-range PCR, however, the orientation could be determined unambiguously, as discussed below.

Long-range PCR to determine orientation of I-CeuI fragments on the genome map. Orientations of I-CeuI fragments B, D, E, F, and G can usually be readily determined by their neighboring relationships with I-CeuI C because the directions of transcription point away from *oriC*, which is located near the *rnmC* end of I-CeuI C (18). However, I-CeuI A, containing *terC*, and I-CeuI C, containing *oriC*, can take either of the two possible orientations. Using 14 primers (Table 1), each for a genomic sequence upstream or downstream of one of the seven *rnm* operons, we determined the orientations of all seven I-CeuI fragments by long-range PCR (Fig. 2). The presence or absence of PCR products with any pairs of primers (Fig. 2A) unambiguously demonstrates the orientations of individual fragments and neighboring relationships of the fragments (Fig.

FIG. 1. Mapping of serovar Gallinarum SARB21 with I-CeuI. (A) PFGE patterns of I-CeuI-cleaved genomic DNA. Lanes: 1, serovar Gallinarum SARB21; 2, λ DNA concatemer as a DNA size marker; 3, serovar Pullorum RKS5078. Frags, fragments. (B) I-CeuI maps of serovar Gallinarum SARB21 and serovar Pullorum RKS5078 based on data shown in panel A. The map of serovar Pullorum RKS5078 is shown here for comparison with that of serovar Gallinarum SARB21.

2B). We found that I-CeuI A in serovar Gallinarum SARB21 was inverted relative to serovar Typhimurium LT2 and that I-CeuI B, D, E, F, and G had the orientations predicted based on the PFGE data. The situation with I-CeuI C was rather surprising. In all previous studies with other *Salmonella* serovars, the *rmC* end of I-CeuI C, thus *oriC* that is in I-CeuI C and

18 kb from the *rmC* end of I-CeuI C (18), is adjacent to at least one of the three small I-CeuI fragments, D, E, and F. We thus had predicted that I-CeuI C might be oriented as in other *Salmonella* serovars, with *oriC* being close to I-CeuI F rather than to I-CeuI G. Figure 2B, however, clearly shows the opposite, which was evidenced by the presence of PCR products

FIG. 2. Results of long-range PCR to determine orientations of the I-CeuI fragments. (A) Electrophoresis gel showing the PCR products. Lane 1 is HindIII-digested λ DNA as a DNA fragment size marker; the other lanes are labeled with the primer pairs used. (B) Orientations of the I-CeuI fragments, as indicated by the pairs of primers (small solid arrows) that gave rise to PCR products. Transcription directions of the *rm* operons are indicated by large hollow arrows. See the text for an explanation of the naming of *rm* hybrids.

FIG. 3. *Tn10* insertion to locate genes. (A) PFGE patterns of *Xba*I-cleaved genomic DNA of serovar *Gallinarum* SARB21 and representative *Tn10* insertion mutants. Lanes: 1, SARB21 with a *Tn10* insertion in *cheA*, leading to the disappearance of fragment A (970 kb) and the appearance of two new fragments, one of 600 kb and one of 380 kb; 2, SARB21 with a *Tn10* insertion in *leuD*, leading to the disappearance of fragment B (620 kb) and the appearance of two new fragments, one of 420 kb and one of 210 kb; 3, wild-type SARB21 strain; 4, λ DNA concatemer as a DNA size marker. Frags, fragments. (B) PFGE patterns of *Avr*II-cleaved genomic DNA of serovar *Gallinarum* SARB21 and representative *Tn10* insertion mutants. Lanes: 1, SARB21 with a *Tn10* insertion in *cheA*, leading to the disappearance of fragment C (522 kb) and the appearance of two new fragments, one of 435 kb and one of 85 kb; 2, SARB21 with a *Tn10* insertion in *leuD*, leading to the disappearance of fragment F (302 kb) and the appearance of two new fragments, one of 165 kb and one of 140 kb; 3, wild-type SARB21 strain; 4, λ DNA concatemer as a DNA size marker. Some smaller fragments ran out of the gel.

with primer pair 12 and 5 (rather than 12 and 11) for the *rmD* end of I-CeuI C and primer pair 14 and 11 (rather than 14 and 5) for the *rmC* end and was further confirmed by *Tn10* insertion analysis (e.g., *uncA* in *Xba*I D, not in *Xba*I O, or *cysG* in *Xba*I F, not in *Xba*I D; see below). These genomic rearrangements led to the *rm* hybrids *rmD/B* and *rrlC/E*. We designated the hybrids in their transcription order, rather than the clockwise genomic order, and thus the *rm* operon between I-CeuI F and C is designated *rmD/B*, not *rmB/D* (Fig. 2B).

Cleavages of serovar *Gallinarum* SARB21 genomic DNA with *Xba*I and *Avr*II and *Tn10* insertions to locate genes. *Xba*I and *Avr*II cleaved the genomic DNA of serovar *Gallinarum*

SARB21 into 18 and 24 fragments, respectively (Fig. 3). To map these fragments and locate genes in the genome, we used the *Tn10* insertion technique. *Tn10* has cleavage sites for both *Xba*I and *Avr*II, so insertion of *Tn10* into known genes will locate those genes (for details of such analysis, see reference 23). By summarizing the results of *Xba*I and *Avr*II digestions of genomic DNA in the *Tn10* mutants, we determined the positions of 60 genes. Double digestions of the cleavage fragments with all combinations, e.g., I-CeuI fragments redigested with *Xba*I or *Avr*II and vice versa (23), enabled us to significantly refine our knowledge of the sizes and locations of the fragments, leading to a physical map of serovar *Gallinarum*

In serovar Typhimurium LT2, I-CeuI fragment B is followed by CDEFG; in serovar Gallinarum SARB21, I-CeuI fragment B is followed by DEFCG. There thus appears to be a simple translocation, moving I-CeuI C to a new location between I-CeuI fragments F and G, which is in fact not the case. Instead, there must have been at least three separate translocations, moving I-CeuI fragments D, E, and F one by one, as judged by the orientations of the fragments D, E, and F (Fig. 2). When I-CeuI fragments D, E, and F are in the serovar Typhimurium order, the *rnn* operon between I-CeuI fragments D and E should be *rnnA* and that between E and F should be *rnnB* (18, 23); for serovar Gallinarum SARB21, the long-range PCR results unambiguously indicate that *rnnB/C* is located between I-CeuI D and E and *rnnE/A* between I-CeuI E and F.

TABLE 2. Genomic segments present in serovar Typhimurium LT2 but absent from serovar Gallinarum SARB21

Region ^a	Size (kb)	BLAST search profile	Identity or product
<i>nadB-purG</i>	46	STM2584–STM2636	Gifsy-1
<i>pyrD-pncB</i>	47	STM1005–STM1056	Gifsy-2
<i>bioA-aspC</i>	57	STM893–STM929	Fels-1
<i>tyrA-cysC</i>	62	STM2694–STM2772	Fels-2 and an island adjacent to Fels-2
<i>recB-pepP</i>	24	STM3022–STM3034	<i>std</i> fimbriae
<i>purA-serB</i>	43	STM4417–STM4437; STM4488–STM4498; STM4520–STM4527; STM4571–STM4575	<i>srjJ</i> ; <i>stj</i> fimbriae; type I restriction system

^a Not inclusive.

The comparisons also revealed two regions with increased lengths and six regions with reduced lengths, which we assumed to be genomic insertions and deletions, respectively. The 31-kb increase between *dadX* and *cheB* is very similar to that in serovar Pullorum RKS5078 both in size and genomic location, whereas the 18-kb increase between *cysC* and *relA* is similar in size to the one in serovar Pullorum RKS5078 between *proU* and *cysC* (19). The six regions with reduced physical lengths include, in the clockwise genomic order, 46 kb between *nadB* and *purG*, 47 kb between *pyrD* and *pncB*, 57 kb between *aspC* and *bioA*, 62 kb between *tyrA* and *cysC*, 24 kb between *recB* and *pepP*, and 43 kb between *purA* and *serB*.

Characterization of putative deletions in serovar Gallinarum SARB21. A strain of serovar Gallinarum, 287/91, is being sequenced at the Sanger Center. Although the sequencing is not finished yet, the Sanger Center provides the database of shotgun reads for BLAST searching. Additionally, Chan et al. (2) have compared genomes of 26 *Salmonella* strains, including serovar Typhimurium LT2, serovar Pullorum SARB51, and serovar Gallinarum SARB21, using serovar Typhimurium microarray. By personal communications with Chan, we obtained the data for genes missing in serovar Gallinarum and serovar Pullorum relative to serovar Typhimurium. Combining information from physical mapping with the microarray data, we selected certain genes for BLAST searching against the database of the Sanger Center for the serovar Gallinarum shotgun reads.

As shown in Table 2, the six genomic regions with reduced physical lengths described above corresponded well with the microarray and BLAST searching results, confirming the absence of genetic material in those regions. The first four regions with reduced physical lengths matched the four prophages in serovar Typhimurium LT2, corresponding in genomic order to Gifsy-1 (STM2584 to STM2636), Gifsy-2 (STM1005 to STM1056), Fels-1 (STM893 to STM929), and Fels-2 (STM2694 to STM2772). (The genes of serovar Typhimurium LT2 are numbered consecutively on the genome from STM1 [35].) Therefore, instead of saying that serovar Gallinarum has these four deletions, it would be better to state that serovar Gallinarum does not have these four prophages.

In addition, we detected the absence of genes equivalent to regions STM3022 to STM3034 (*recB-pepP*), STM4417 to STM4437, STM4488 to STM4498, STM4520 to STM4527, and STM4571 to STM4575 (*purA-serB*), not inclusive, in serovar Gallinarum. Of these areas, STM3022 to STM3034 contains genes encoding *std* fimbriae, STM4417 to STM4436 contains *srjJ* genes, STM4571 to STM4575 encodes *stj* fimbriae,

and genes in STM4520 to STM4527 code for a type I restriction system (35).

Genomic comparisons between serovars Gallinarum and Pullorum. Upon comparing the genome maps of serovar Gallinarum SARB21 and serovar Pullorum RKS5078, which was published previously (19), we found that they were similar in many ways. Firstly, both strains seemed to have identical inversions 1 and 2, although inversion 2 in serovar Pullorum RKS5078 has both endpoints within, or close to, genomic insertions, leaving a question as to whether the two bacteria have the same endpoints. Secondly, the two putative insertions in serovar Gallinarum SARB21 seemed to have counterparts in serovar Pullorum RKS5078 at similar genomic locations, although their homology needs to be examined. Finally, the two strains have in common the absence of several DNA segments relative to serovar Typhimurium, including the four prophages.

The differences between the two genomic maps are also significant. Among them are differences in genomic sizes, with the genome of serovar Gallinarum SARB21 being as much as 190 kb smaller. The other differences include individual genomic insertions or deletions found in one or the other strain. For example, between *purA* and *serB* on the genome map of serovar Gallinarum SARB21, the length is 43 kb shorter than that of the corresponding region in serovar Typhimurium LT2, while in the corresponding region of serovar Pullorum RKS5078, there is a 20-kb decrease between *purA* and *pyrB* and a 12-kb increase between *argI* and *serB*. Significantly, the 157-kb insertion in serovar Pullorum RKS5078 between *pncX* and *pyrF* was not found in serovar Gallinarum SARB21, although we do not exclude the possibility that parts of it may be distributed in the genome of serovar Gallinarum. Difference in orientations of I-CeuI C is another genomic feature to distinguish the two strains.

DISCUSSION

All *Salmonella* serovars share a general gene pool for essential cell functions, and each has a sub-gene pool for adaptation to a specific niche. Genetic barriers seem to exist among the *Salmonella* lineages, as subtle genomic features are shared only by members of the same lineage (a phylogenetic group, such as a biovar, serovar, species, and etc.) and younger serovars such as serovar Typhi tend to be more homogenous than the more ancient serovars such as serovar Typhimurium (4, 13, 28, 30, 38). As a result, at least for *Salmonella* and especially for the host-adapted and thus younger lineages, a bacterial strain should essentially reflect one lineage in gene content. Based on

this assumption, which has been supported by previous work, we have constructed genome maps for representative *Salmonella* serovars. In this study, we mapped the genome of serovar Gallinarum SARB21 and compared it with genomes of other *Salmonella* serovars. We identified several genomic differences between serovar Gallinarum SARB21 and serovar Typhimurium LT2 and between serovar Gallinarum SARB21 and serovar Pullorum RKS5078, including genomic increases, decreases, and rearrangements relative to one or the other *Salmonella* lineage. While attempting to isolate the putative insertions by the suppression subtractive hybridization techniques for genetic and functional studies (S.-L. Liu, ongoing work), we analyzed the rearrangements (inversions and translocations) by using long-range PCR and characterized the genomic regions that have reduced physical lengths compared with corresponding regions in serovar Typhimurium LT2 by using bioinformatics tools.

Among the several kinds of genomic rearrangements in serovar Gallinarum SARB21, one was rare and rather surprising, i.e., the orientation of I-CeuI C with regard to I-CeuI DEF. In all of the hundreds of previously analyzed *Salmonella* strains, I-CeuI C is oriented so that *oriC*, which is in I-CeuI C and 18 kb from the *rmC* end of I-CeuI C, is close to at least one of the three small I-CeuI fragments, D, E and F. In serovar Gallinarum SARB21, however, the I-CeuI C fragment has its *rmC* end (and hence *oriC*) away from all of I-CeuI DEF. In the adopt-adapt model of bacterial evolution in exploration of mechanisms for genomic rearrangements, rearranging for rebalancing of the genome has been emphasized, with dosage effect as an additional or alternative explanation (19, 26, 30, 31). The case of I-CeuI C orientation in serovar Gallinarum SARB21 seems to provide stronger support to the rearranging for rebalancing hypothesis than to the dosage effect alternative, as all genes on the I-CeuI D, E, and F fragments will have relatively reduced dosage when they are all so much farther away from *oriC*, although we need to confirm this by obtaining further evidence, such as the balance status in this strain.

Another point of great interest in the comparison of fowl-adapted *Salmonella* serovars is the fact that both serovar Gallinarum SARB21 and serovar Pullorum RKS5078 have genomic rearrangements in the genomic region between *rmD* and *rmE*. In a recent study of pigeon-associated serovar Typhimurium DT2 strains, Helm et al. also detected genomic rearrangements in the *rmD*- and *rmE*-flanked region (9). Whether rearrangements in this region would relate to bird adaptation of the *Salmonella* pathogens remains to be investigated.

Recombination involving *rm* genes, resulting in tandem duplications or other rearrangements, has been an active area of study in regard to its significance in bacterial life activities, fitness, and evolution (7, 10–12, 15, 26, 30, 42). Previous studies have explored the reasons why some bacteria undergo frequent recombination events but many of their close relatives do not, such as in the case of serovar Typhi versus serovar Typhimurium (26, 28, 30). Previously, it was demonstrated that bacterial genomes have a tendency to obtain a physical balance (44). Based on this and other evidence, it was hypothesized that recombination may help the bacteria to rebalance the genome following major genomic insertions or deletions, which may have disrupted the genomic balance (19, 30, 31), although this hypothesis needs to be tested further as excep-

tions awaiting explanation also exist. Genomic rebalancing involves homologous recombination, and *rm* genes provide ideal sites for this because of their large sizes and multiple copies (12, 14, 15, 21), although homologous recombination occurs also in other multiple-copy sequences, such as IS200 (1).

Among the genomic regions with reduced physical lengths in serovar Gallinarum SARB21 compared to those in serovar Typhimurium LT2 are four locations where there are prophages in serovar Typhimurium LT2 but not in serovar Gallinarum SARB21. Microarray analyses demonstrate that these prophages are limited to only some strains of serovar Typhimurium, not widespread among *Salmonella* serovars (41). Other genomic regions with reduced physical lengths in serovar Gallinarum SARB21 contain several small deletions. Interestingly, the missing genes in those locations in serovar Gallinarum SARB21 are clustered within islands in serovar Typhimurium LT2. Most of them are adjacent to a tRNA and encode transposase or integrase, suggesting that these islands are mobile and may have been acquired by horizontal gene transfer (35, 37).

The overall genomic similarity between serovars Gallinarum and Pullorum, including those genomic features common only to these two *Salmonella* lineages, further indicates their close relatedness, in addition to the previously known facts that they are both nonflagellate, specific to chickens, of the same antigenic formula, and hard to distinguish biochemically. With genomic sequence shotgun reads of a strain of serovar Gallinarum available from the Sanger Center for BLAST searching, we are working with Rob Edwards of Tennessee University to sequence a strain of serovar Pullorum, now at the ca. 80% genomic coverage stage. The genomic increases and decreases revealed by physical mapping of the genomes of serovars Gallinarum and Pullorum relative to the genome of serovar Typhimurium LT2 are guiding us to focus on regions of interest to identify the common loss of ancestral genes that may account for the fowl specificity of these pathogens and their unique genetic materials that may be responsible for their distinct pathogenesises, potentially leading to new insights into the genomic evolution and speciation of bacteria and the emergence of new bacterial pathogens.

ACKNOWLEDGMENTS

We thank Ning Qi, Lili Lei, Sushma Kothapalli, and Suneetha Alokam for technical assistance.

The work was supported by Discovery grants from the Natural Sciences and Engineering Research Council (NSERC) to S.-L.L. and K.E.S., by a grant from the National Institute of Allergy and Infectious Diseases of the National Institutes of Health to K.E.S. (AI-034829-14), and by a grant from the Canadian Institute of Health Research (CIHR) to R.N.J. W.-Q.L. was supported by a summer studentship from Alberta Heritage Foundation for Medical Research, and A.Q.W. and S.Z. were both supported by summer studentships from NSERC.

REFERENCES

1. Alokam, S., S. L. Liu, K. Said, and K. E. Sanderson. 2002. Inversions over the terminus region in *Salmonella* and *Escherichia coli*: IS200s as the sites of homologous recombination inverting the chromosome of *Salmonella enterica* serovar Typhi. *J. Bacteriol.* **184**:6190–6197.
2. Chan, K., S. Baker, C. C. Kim, C. S. Detweiler, G. Dougan, and S. Falkow. 2003. Genomic comparison of *Salmonella enterica* serovars and *Salmonella bongori* by use of an *S. enterica* serovar Typhimurium DNA microarray. *J. Bacteriol.* **185**:553–563.
3. Crosa, J. H., D. J. Brenner, W. H. Ewing, and S. Falkow. 1973. Molecular relationships among the *Salmonellae*. *J. Bacteriol.* **115**:307–315.

4. Deng, W., S. R. Liou, G. Plunkett III, G. F. Mayhew, D. J. Rose, V. Burland, V. Kodoyianni, D. C. Schwartz, and F. R. Blattner. 2003. Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2 and CT18. *J. Bacteriol.* **185**:2330–2337.
5. Edwards, R. A., A. G. J. Olsen, and S. R. Maloy. 2002. Comparative genomics of closely related salmonellae. *Trends Microbiol.* **10**:94–99.
6. Gauthier, A., M. Turmel, and C. Lemieux. 1991. A group I intron in the chloroplast large subunit rRNA gene of *Chlamydomonas eugametos* encodes a double-strand endonuclease that cleaves the homing site of this intron. *Curr. Genet.* **19**:43–47.
7. Helm, R. A., A. G. Lee, H. D. Christman, and S. Maloy. 2003. Genomic rearrangements at *rrn* operons in *Salmonella*. *Genetics* **165**:951–959.
8. Helm, R. A., and S. Maloy. 2001. Rapid approach to determine *rm* arrangement in *Salmonella* serovars. *Appl. Environ. Microbiol.* **67**:3295–3298.
9. Helm, R. A., S. Porwollik, A. E. Stanley, S. Maloy, M. McClelland, W. Rabusch, and A. Eisenstark. 2004. Pigeon-associated strains of *Salmonella enterica* serovar Typhimurium phage type DT2 have genomic rearrangements at rRNA operons. *Infect. Immun.* **72**:7338–7341.
10. Hill, C. W., R. H. Grafstrom, B. W. Harnish, and B. S. Hillman. 1977. Tandem duplications resulting from recombination between ribosomal RNA genes in *Escherichia coli*. *J. Mol. Biol.* **116**:407–428.
11. Hill, C. W., and J. A. Gray. 1988. Effects of chromosomal inversion on cell fitness in *Escherichia coli* K-12. *Genetics* **119**:771–778.
12. Hill, C. W., and B. W. Harnish. 1981. Inversions between ribosomal RNA genes of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **78**:7069–7072.
13. Kidgell, C., U. Reichard, J. Wain, B. Linz, M. Torpdahl, G. Dougan, and M. Achtman. 2002. *Salmonella typhi*, the causative agent of typhoid fever, is approximately 50,000 years old. *Infect. Genet. Evol.* **2**:39–45.
14. Kothapalli, S., S. Nair, S. Alokam, T. Pang, R. Khakhria, D. Woodward, W. Johnson, B. A. Stocker, K. E. Sanderson, and S. L. Liu. 2005. Diversity of genome structure in *Salmonella enterica* serovar Typhi populations. *J. Bacteriol.* **187**:2638–2650.
15. Lehner, A. F., and C. W. Hill. 1980. Involvement of ribosomal ribonucleic acid operons in *Salmonella typhimurium* chromosomal rearrangements. *J. Bacteriol.* **143**:492–498.
16. Le Minor, L. 1988. Typing of *Salmonella* species. *Eur. J. Clin. Microbiol. Infect. Dis.* **7**:214–218.
17. Li, J., N. H. Smith, K. Nelson, P. B. Crichton, D. C. Old, T. S. Whittam, and R. K. Selander. 1993. Evolutionary origin and radiation of the avian-adapted non-motile salmonellae. *J. Med. Microbiol.* **38**:129–139.
18. Liu, G. R., K. Edwards, A. Eisenstark, Y. M. Fu, W. Q. Liu, K. E. Sanderson, R. N. Johnston, and S. L. Liu. 2003. Genomic diversification among archival strains of *Salmonella enterica* serovar Typhimurium LT7. *J. Bacteriol.* **185**:2131–2142.
19. Liu, G. R., A. Rahn, W. Q. Liu, K. E. Sanderson, R. N. Johnston, and S. L. Liu. 2002. The evolving genome of *Salmonella enterica* serovar Pullorum. *J. Bacteriol.* **184**:2626–2633.
20. Liu, S. L., A. Hessel, H. Y. Cheng, and K. E. Sanderson. 1994. The *XbaI-BlnI-CeuI* genomic cleavage map of *Salmonella paratyphi* B. *J. Bacteriol.* **176**:1014–1024.
21. Liu, S. L., A. Hessel, and K. E. Sanderson. 1993. Genomic mapping with I-Ceu I, an intron-encoded endonuclease specific for genes for ribosomal RNA, in *Salmonella* spp., *Escherichia coli*, and other bacteria. *Proc. Natl. Acad. Sci. USA* **90**:6874–6878.
22. Liu, S. L., A. Hessel, and K. E. Sanderson. 1993. The *XbaI-BlnI-CeuI* genomic cleavage map of *Salmonella enteritidis* shows an inversion relative to *Salmonella typhimurium* LT2. *Mol. Microbiol.* **10**:655–664.
23. Liu, S. L., A. Hessel, and K. E. Sanderson. 1993. The *XbaI-BlnI-CeuI* genomic cleavage map of *Salmonella typhimurium* LT2 determined by double digestion, end labelling, and pulsed-field gel electrophoresis. *J. Bacteriol.* **175**:4104–4120.
24. Liu, S. L., and K. E. Sanderson. 1995. The chromosome of *Salmonella paratyphi* A is inverted by recombination between *rmH* and *rmG*. *J. Bacteriol.* **177**:6585–6592.
25. Liu, S. L., and K. E. Sanderson. 1995. Genomic cleavage map of *Salmonella typhi* Ty2. *J. Bacteriol.* **177**:5099–5107.
26. Liu, S. L., and K. E. Sanderson. 1996. Highly plastic chromosomal organization in *Salmonella typhi*. *Proc. Natl. Acad. Sci. USA* **93**:10303–10308.
27. Liu, S. L., and K. E. Sanderson. 1998. Homologous recombination between *rrn* operons rearranges the chromosome in host-specialized species of *Salmonella*. *FEMS Microbiol. Lett.* **164**:275–281.
28. Liu, S. L., and K. E. Sanderson. 1995. I-CeuI reveals conservation of the genome of independent strains of *Salmonella typhimurium*. *J. Bacteriol.* **177**:3355–3357.
29. Liu, S. L., and K. E. Sanderson. 1992. A physical map of the *Salmonella typhimurium* LT2 genome made by using *XbaI* analysis. *J. Bacteriol.* **174**:1662–1672.
30. Liu, S. L., and K. E. Sanderson. 1995. Rearrangements in the genome of the bacterium *Salmonella typhi*. *Proc. Natl. Acad. Sci. USA* **92**:1018–1022.
31. Liu, S. L., A. B. Schryvers, K. E. Sanderson, and R. N. Johnston. 1999. Bacterial phylogenetic clusters revealed by genome structure. *J. Bacteriol.* **181**:6747–6755.
32. Marshall, P., T. B. Davis, and C. Lemieux. 1994. The I-CeuI endonuclease: purification and potential role in the evolution of *Chlamydomonas* group I introns. *Eur. J. Biochem.* **220**:855–859.
33. Marshall, P., and C. Lemieux. 1991. Cleavage pattern of the homing endonuclease encoded by the fifth intron in the chloroplast large subunit rRNA-encoding gene of *Chlamydomonas eugametos*. *Gene* **104**:241–245.
34. McClelland, M., K. E. Sanderson, S. W. Clifton, P. Latreille, S. Porwollik, A. Sabo, R. Meyer, T. Bieri, P. Ozersky, M. McLellan, C. R. Harkins, C. Wang, C. Nguyen, A. Berghoff, G. Elliott, S. Kohlberg, C. Strong, F. Du, J. Carter, C. Kremizki, D. Layman, S. Leonard, H. Sun, L. Fulton, W. Nash, T. Miner, P. Minx, K. Delehaunty, C. Fronick, V. Magrini, M. Nhan, W. Warren, L. Florea, J. Spieth, and R. K. Wilson. 2004. Comparison of genome degradation in *Paratyphi* A and Typhi, human-restricted serovars of *Salmonella enterica* that cause typhoid. *Nat. Genet.* **36**:1268–1274.
35. McClelland, M., K. E. Sanderson, J. Spieth, S. W. Clifton, P. Latreille, L. Courtney, S. Porwollik, J. Ali, M. Dante, F. Du, S. Hou, D. Layman, S. Leonard, C. Nguyen, K. Scott, A. Holmes, N. Grewal, E. Mulvaney, E. Ryan, H. Sun, L. Florea, W. Miller, T. Stoneking, M. Nhan, R. Waterston, and R. K. Wilson. 2001. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* **413**:852–856.
36. Nair, S., S. Alokam, S. Kothapalli, S. Porwollik, E. Proctor, C. Choy, M. McClelland, S. L. Liu, and K. E. Sanderson. 2004. *Salmonella enterica* serovar Typhi strains from which SPI7, a 134-kilobase island with genes for Vi exopolysaccharide and other functions, has been deleted. *J. Bacteriol.* **186**:3214–3223.
37. Ochman, H., J. G. Lawrence, and E. A. Groisman. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**:299–304.
38. Parkhill, J., G. Dougan, K. D. James, N. R. Thomson, D. Pickard, J. Wain, C. Churcher, K. L. Mungall, S. D. Bentley, M. T. Holden, M. Sebaihia, S. Baker, D. Basham, K. Brooks, T. Chillingworth, P. Connerton, A. Cronin, P. Davis, R. M. Davies, L. Dowd, N. White, J. Farrar, T. Feltwell, N. Hamlin, A. Haque, T. T. Hien, S. Holroyd, K. Jagels, A. Krogh, T. S. Larsen, S. Leather, S. Moule, P. O'Gaora, C. Parry, M. Quail, K. Rutherford, M. Simmonds, J. Skelton, K. Stevens, S. Whitehead, and B. G. Barrell. 2001. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature* **413**:848–852.
39. Popoff, M. Y., J. Bockemuhl, and L. L. Gheesling. 2004. Supplement 2002 (no. 46) to the Kauffmann-White scheme. *Res. Microbiol.* **155**:568–570.
40. Porwollik, S., J. Frye, L. D. Florea, F. Blackmer, and M. McClelland. 2003. A non-redundant microarray of genes for two related bacteria. *Nucleic Acids Res.* **31**:1869–1876.
41. Porwollik, S., and M. McClelland. 2003. Lateral gene transfer in *Salmonella*. *Microbes Infect.* **5**:977–989.
42. Porwollik, S., R. M. Wong, R. A. Helm, K. K. Edwards, M. Calcutt, A. Eisenstark, and M. McClelland. 2004. DNA amplification and rearrangements in archival *Salmonella enterica* serovar Typhimurium LT2 cultures. *J. Bacteriol.* **186**:1678–1682.
43. Shivaprasad, H. L. 2000. Fowl typhoid and pullorum disease. *Rev. Sci. Tech.* **19**:405–424.
44. Song, J., A. Ware, and S. L. Liu. 2003. Wavelet to predict bacterial ori and ter: a tendency towards a physical balance. *BMC Genomics* **4**:17. [Online.] doi:10.1186/1471-2164-4-17.