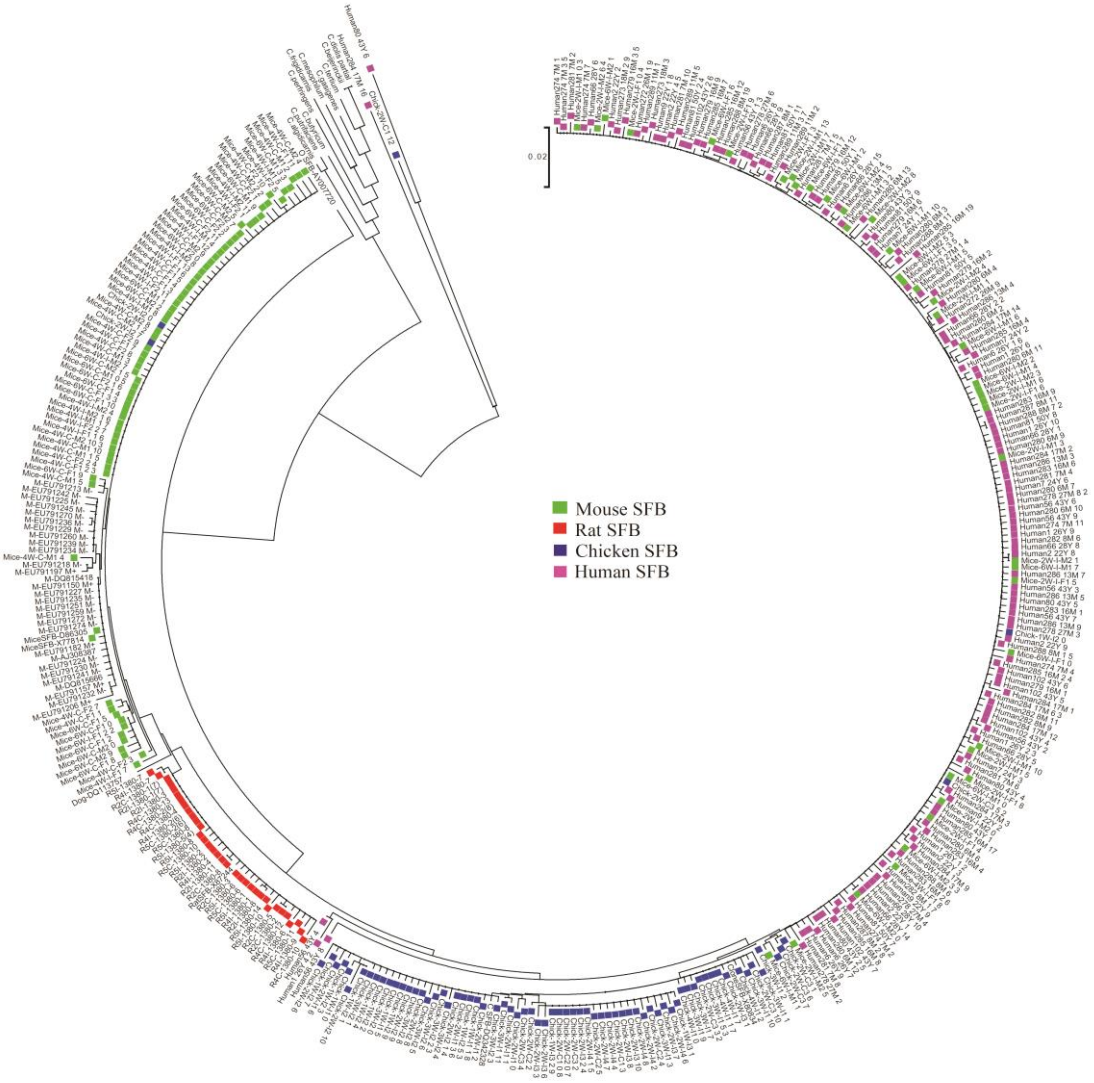
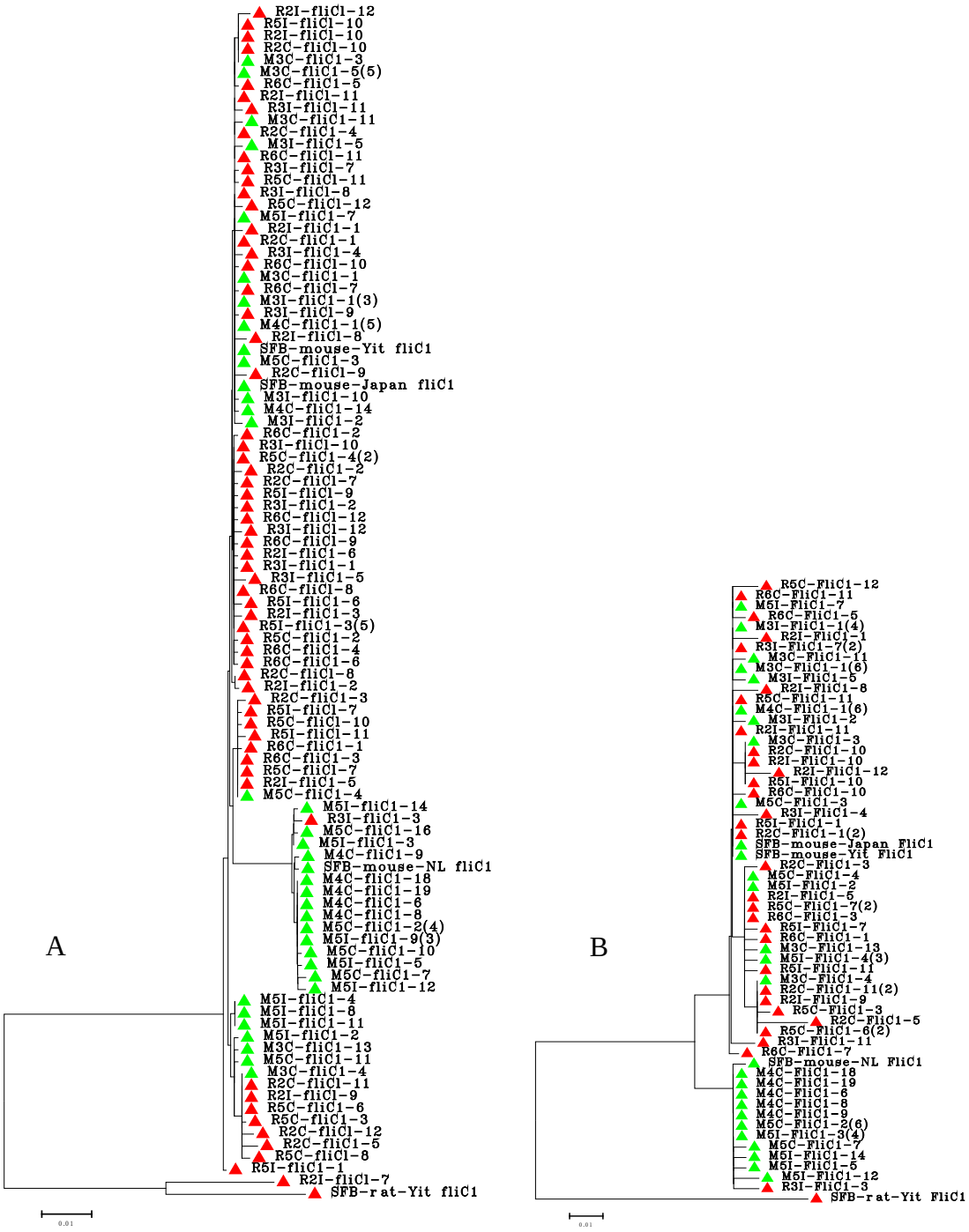
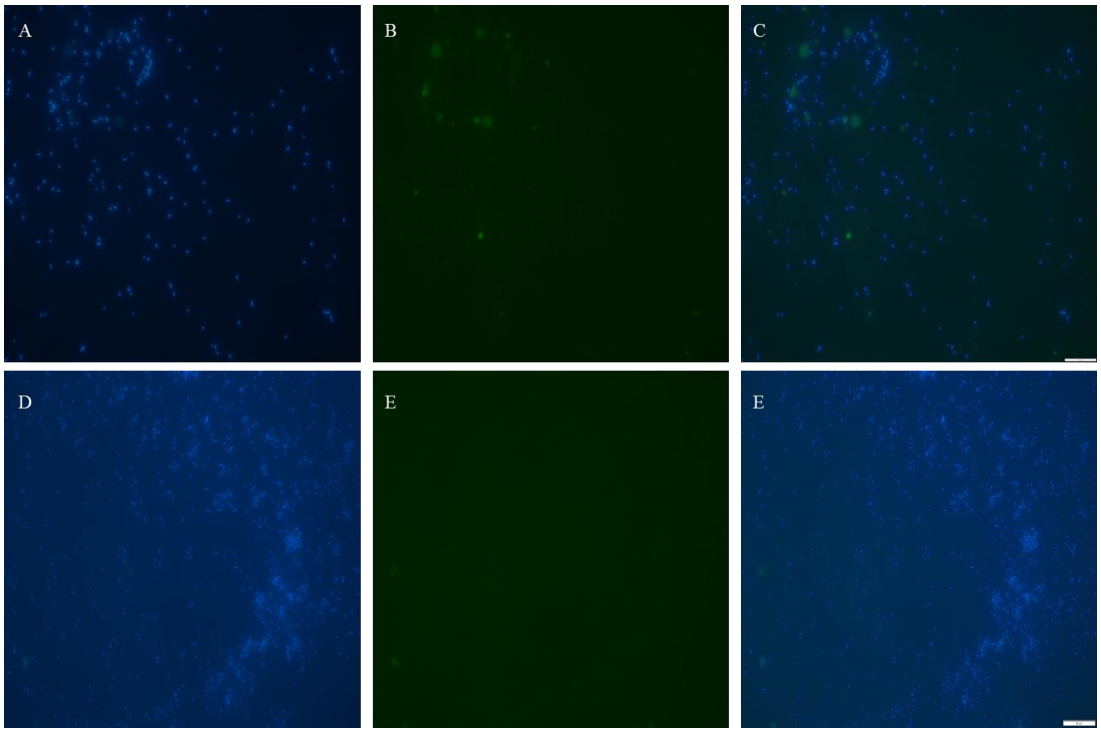
Supplementary Figure 1. Host specificity analysis of SFB 16S rRNA gene sequences.



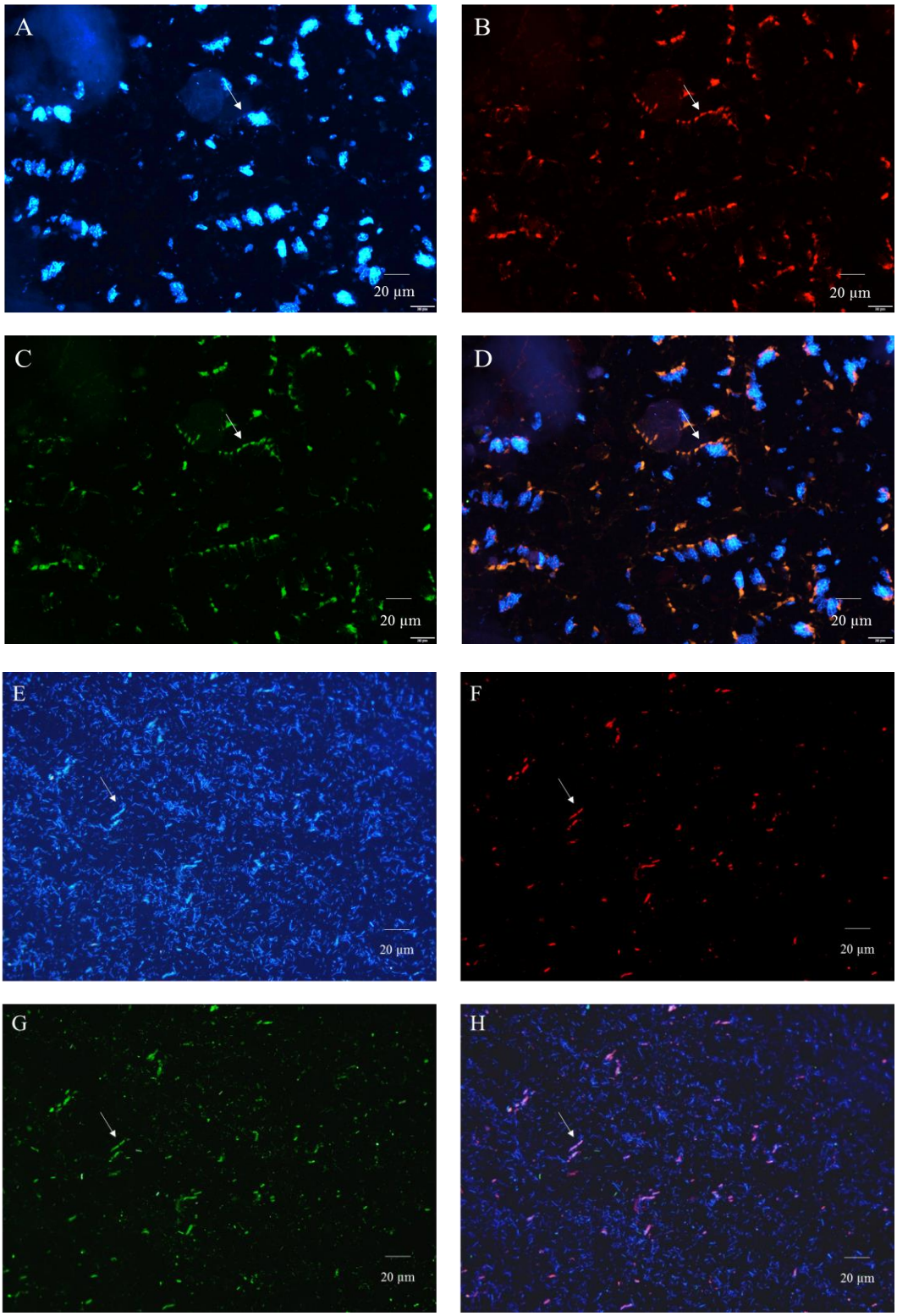
Supplementary Figure 2. Host specificity analysis of SFB *fliC1* at nucleotide and deduced amino acid levels.



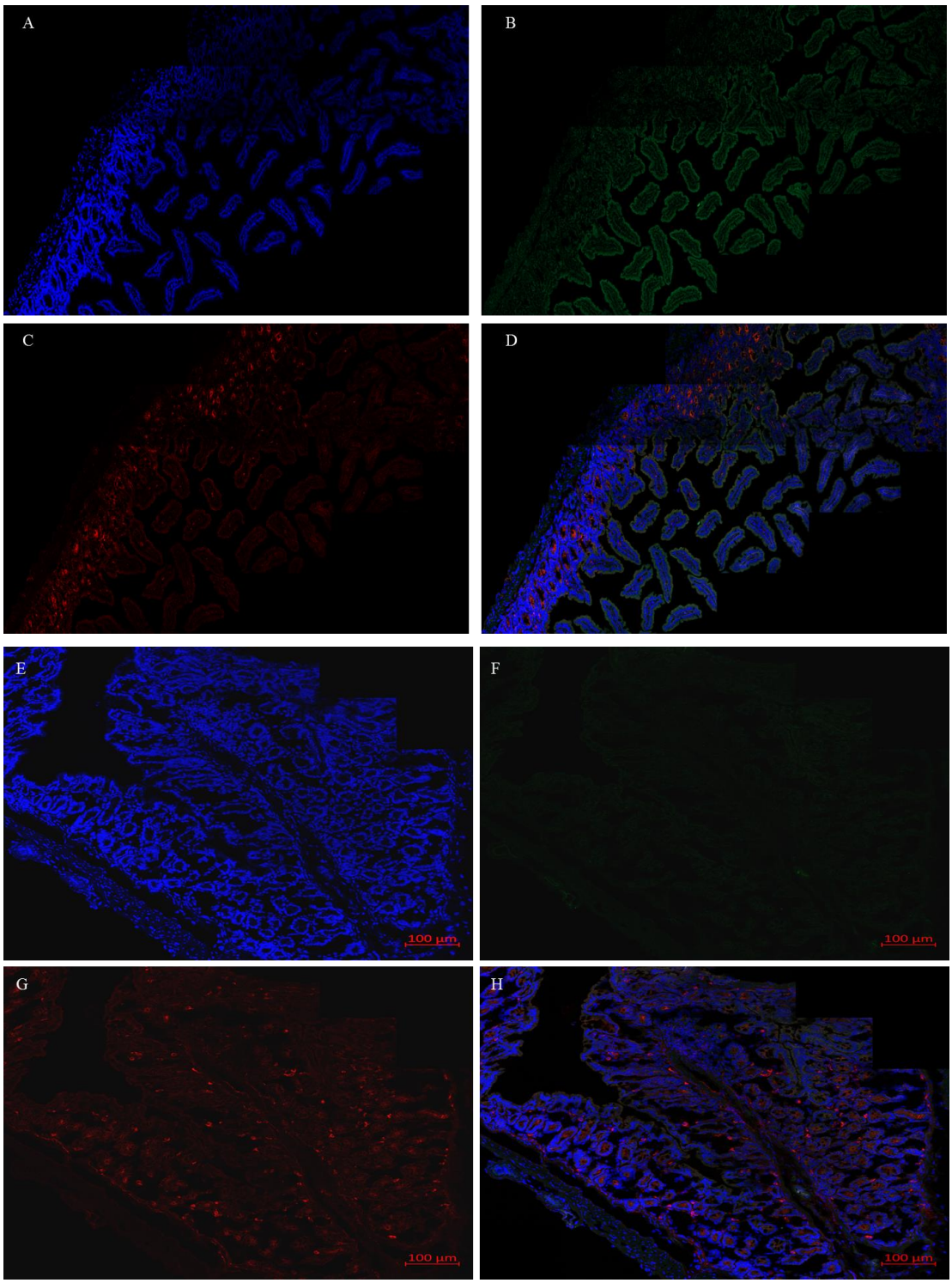
Supplementary Figure 3. Staining of *E. coli* BL21 and *Salmonella* CVCC519 with DAPI and SFB FliC3 antibody.



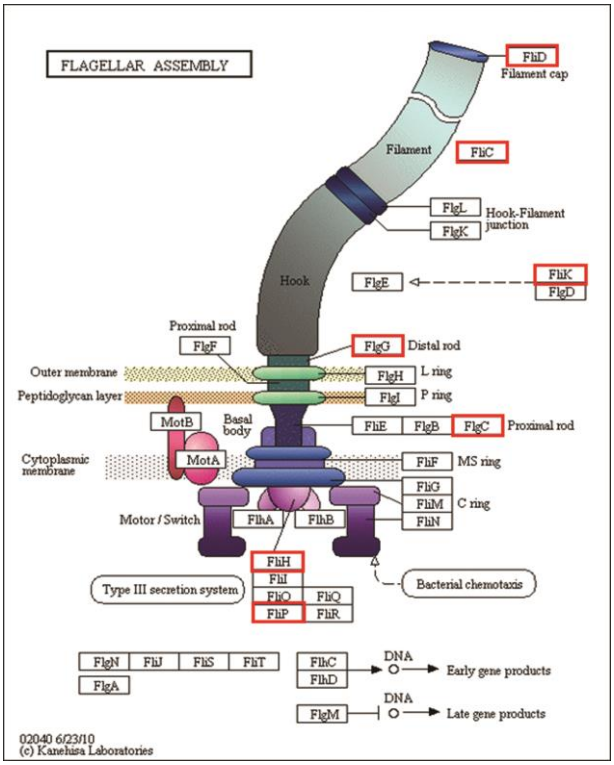
Supplementary Figure 4. Expression of SFB FliC3 in mouse ileum mucosa and cecal content using FISH and IHC analysis.



Supplementary Figure 5. Relative location of SFB in gut epithelial tissue.

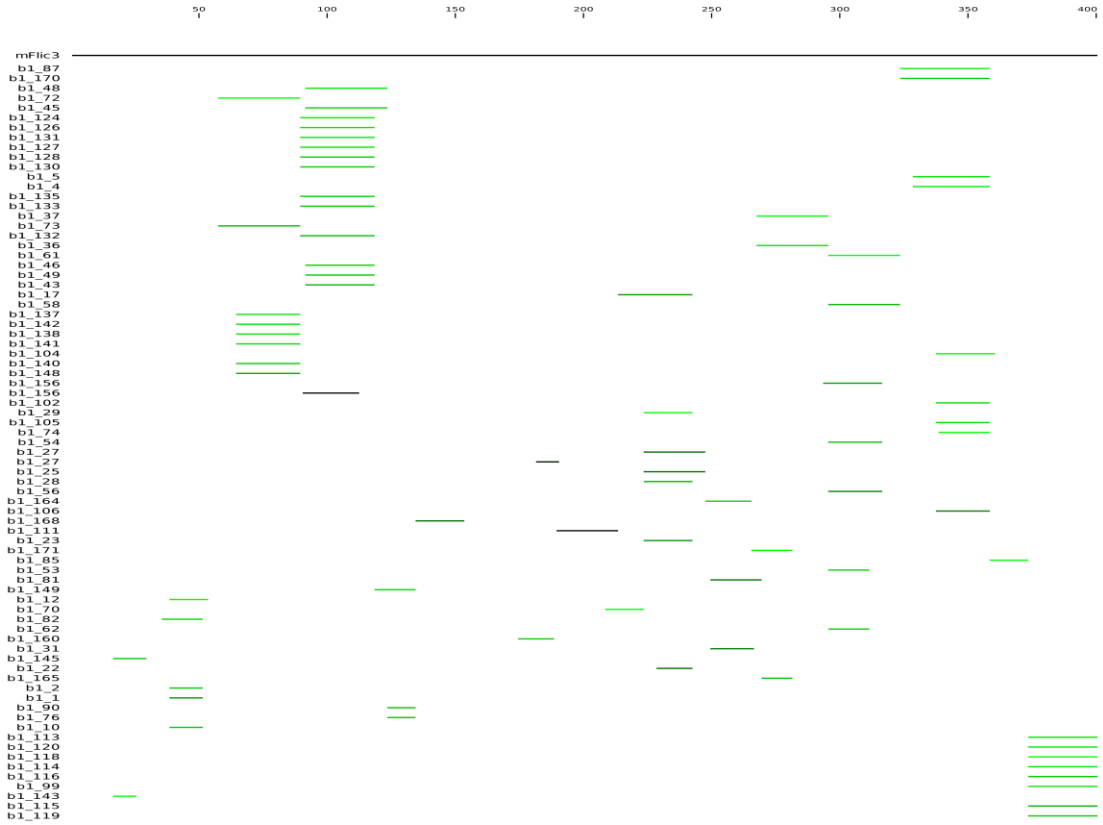


Supplementary Figure 6. Identification of genes related to flagellar assembly using LC/MC/MC in mouse gut samples.

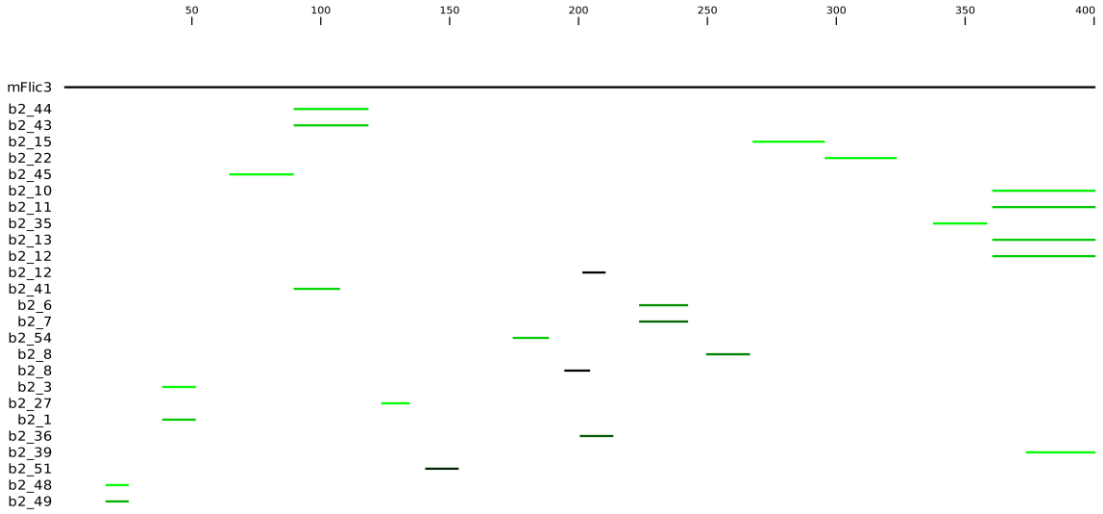


Supplementary Figure 7. Peptides identified from purified mFliC3 bands.

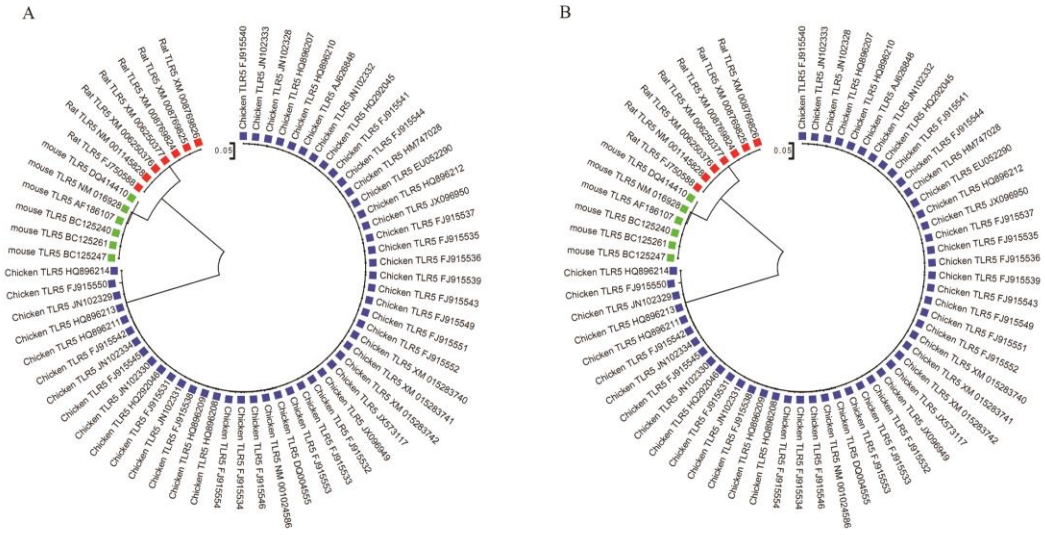
A



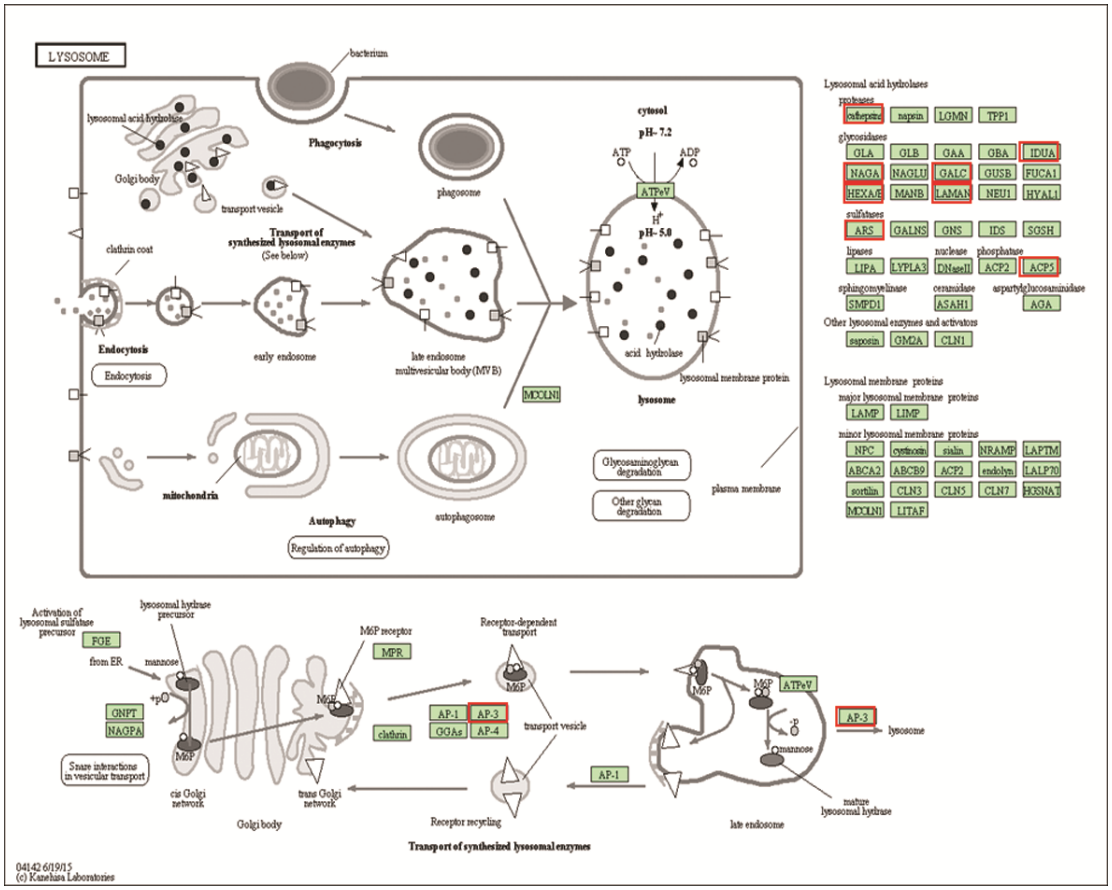
B



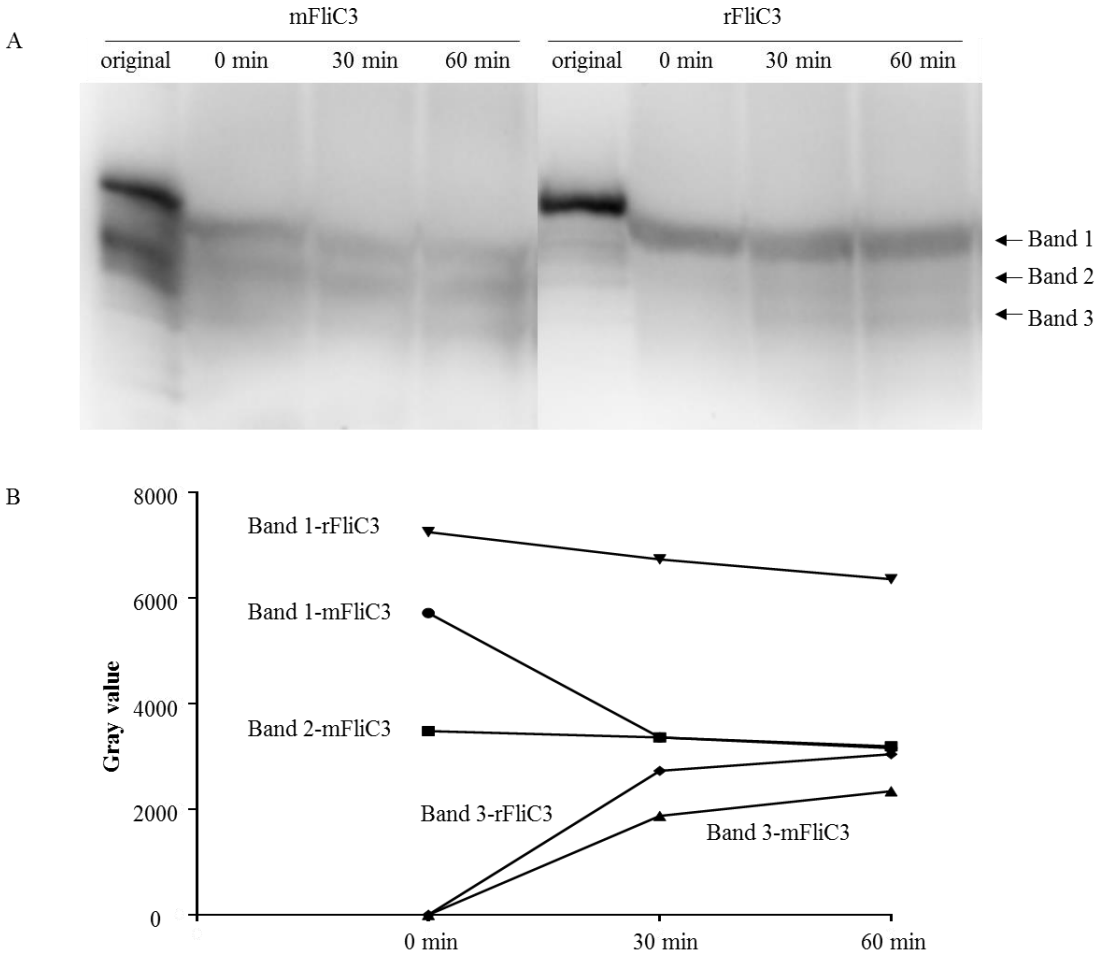
Supplementary Figure 8. Phylogenetic analysis of TLR5 genes at the nucleotide and deduced amino acid levels.



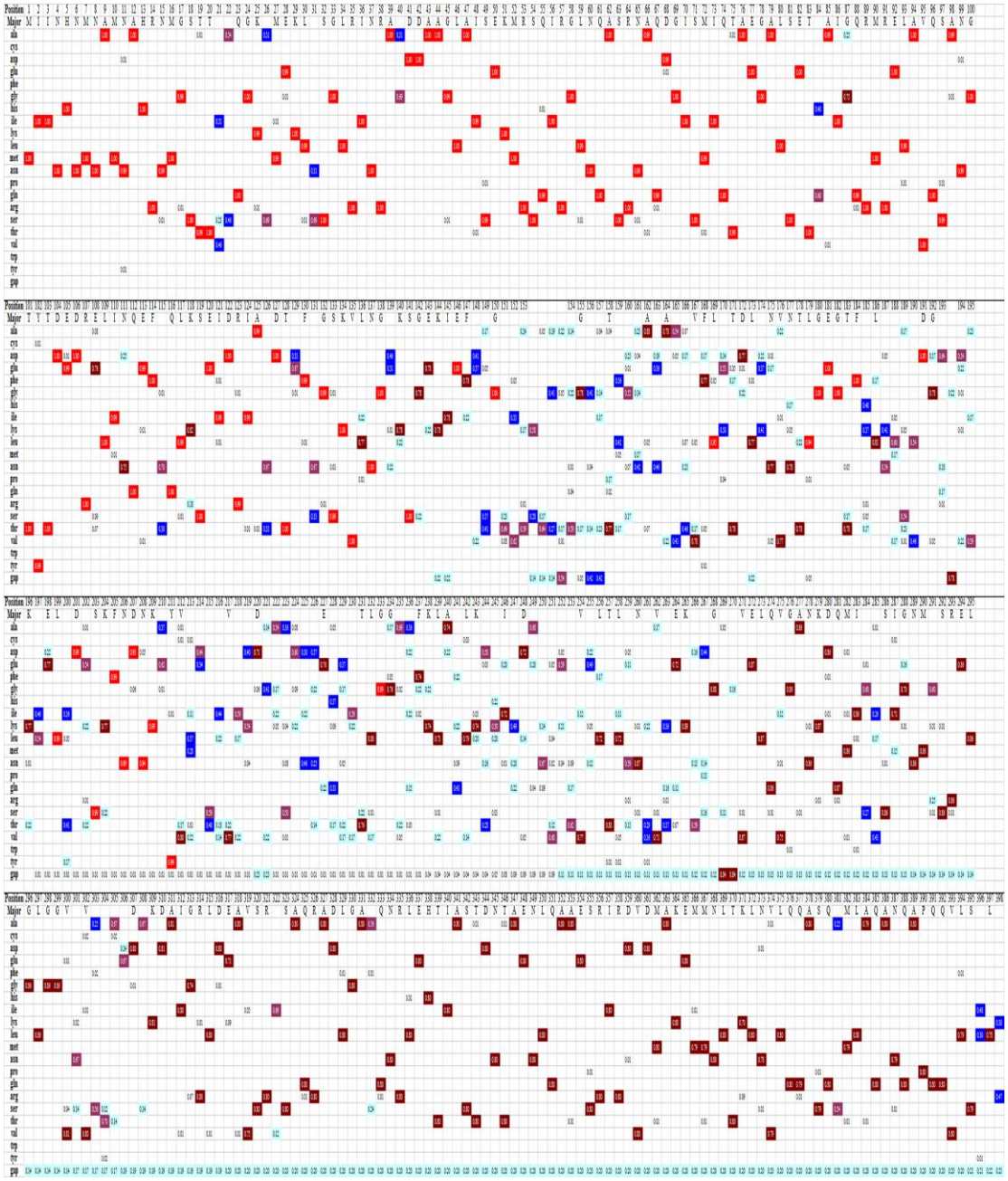
Supplementary Figure 9. RFliC3 enriched proteins related to the lysosome pathway.



Supplementary Figure 10. Degradation of mFliC3 and rFliC3 by rat ileum tissue proteins.



Supplementary Figure 11. Visualization of mutant variable regions of the deduced SFB FliC3 and FliC4 amino acid sequences.



Supplementary Table 1. Summary PCR detection results of SFB genes.

Sample			779F-1380R	*fliC 1*	*fliC 2*	*fliC 3*	*fliC 4*
SFB-Mouse-Japan*			+	+	+	+	+
SFB-Mouse-NYU*			+	–	+	–	+
SFB-Mouse-SU*			+	–	+	–	+
SFB-Mouse-Yit*			+	–	+	–	+
Mouse	1	Ileum	+	+	+	+	+
		Cecum	+	+	+	+	+
Mouse	2	Ileum	+	+	+	+	+
		Cecum	+	+	+	+	+
Mouse	3	Ileum	+	+	+	+	+
		Cecum	+	+	+	+	+
Mouse	4	Ileum	+	+	+	+	+
		Cecum	+	+	+	+	+
Mouse	5	Ileum	+	+	+	+	+
		Cecum	+	+	+	+	+
Mouse	6	Ileum	+	+	+	+	+
		Cecum	+	+	+	+	+
SFB-Rat-Yit*			+	–	+	+	+
Rat	1	Ileum	+	+	–	+	+
		Cecum	+	+	–	+	+
Rat	2	Ileum	+	+	–	+	+
		Cecum	+	+	–	+	+
Rat	3	Ileum	+	+	–	+	+
		Cecum	+	+	–	+	+
Rat	4	Ileum	+	+	–	+	+
		Cecum	+	+	–	+	+
Rat	5	Ileum	+	+	–	+	+
		Cecum	+	+	–	+	+
Rat	6	Ileum	+	+	–	+	+
		Cecum	+	+	–	+	+

*, represents genes identified from full-length genome sequences.

Supplementary Table 2. The diversity of mouse SFB flagellin genes.

Mouse genes	Sample name	No. of sequences	Individual difference		
			Nucleotide acid homology	Amino acid homology	Stop codon
fliC1 840 bp 279 aa	M3C	10	99.2%-100% (5)	98.9-100% (6)	NA
	M3I	6	99.5%-100% (3)	99.3%-100% (4)	NA
	M4C	11	97.8%-100% (5)	97.8-100% (6)	NA
	M5C	10	97.4%-100% (4)	97.1%-100% (6)	58 aa (1)
	M5I	12	97.3%-100% (3)	96.4%-100% (4)	NA
fliC2 1191 bp 396 aa	M3C	10	99.7%-100% (4)	99.2%-100% (5)	NA
	M3I	10	99.6%-100% (2)	89.0%-100% (4)	363 aa (1)
	M4C	10	99.7%-100% (2)	99.2%-100% (2)	NA
	M4I	11	99.2%-99.9%	98.7%-100% (3)	NA
	M5C	10	99.3%-100% (2)	95.3%-100% (2)	271 aa (1), 276 aa (1)
	M5I	10	99.4%-100% (2)	98.7%-100% (2)	298 aa (1), 380 aa (1)
fliC3 1203 bp 400 aa	M3C	11	81.6%-99.7%	76.4%-100% (4)	109 aa (5), 292 aa (1), 315 aa (1), 397 aa (1), 398 aa (3)
	M3I	10	81.8%-100% (2)	81.9%-100% (2)	81 aa (1), 109 aa (2), 397 aa (2), 398 aa (5)
	M4C	10	83.2%-99.8%	82.1%-100% (5)	397 aa (7), 398 aa (3)
	M4I	13	82.4%-99.7%	82.4%-100% (4)	1397 aa (10), 398 aa (3)
	M5C	10	90.4%-100% (2)	81.6%-100% (3)	304 aa (1), 397 aa (8), 398 aa (1)
	M5I	12	83.1%-99.7%	81.9%-100% (3)	397 aa (9), 398 aa (3)
fliC4 1194 bp 397 aa	M3C	10	81.9%-99.9%	82.4%-99.5% (4)	109 aa (5), 300 aa (1), 365 aa (1), 397 aa (1)
	M3I	10	81.9%-99.9%	66.7%-100% (2)	6 aa (1), 9 aa (1), 109 aa (2), 217 aa (1)
	M4C	10	82.2%-99.8%	82.3%-93.7%	104 aa (1), 119 aa (2), 395 aa (1), 397 aa (4)
	M5C	10	83.2%-100% (2)	81.6%-100% (2)	197 aa (1), 304 aa (1)
	M5I	11	83.4%-100% (2)	76.4%-100% (5)	305 aa (1)

NA, not analyzed. Numbers in parentheses represent number of duplicate sequences.

Supplementary Table 3. The diversity of rat SFB genes.

Rat genes	Sample name	No. of sequences	Individual difference		
			Nucleotide acid homology	Amino acid homology	Stop codon
fliC1 840 bp 279 aa	R2C	11	98.5%-100% (2)	97.4%-100% (2)	29 aa (3)
	R5C	10	99.0%-100%	98.1%-100% (2)	29 aa (3)
	R6C	12	99.3%-99.9%	98.9%-99.6%	29 aa (6)
	R2I	11	86.6%-99.9%	98.1%-99.6%	29 aa (3), 61 aa (1)
	R3I	11	97.1%-99.9%	96.3%-100% (2)	29 aa (5)
	R5I	11	99.3%-100% (5)	98.9%-99.3%	29 aa (7)
fliC3 1203 bp 400 aa	R2C	10	75.5%-100% (2)	70.3%-100% (3)	41 aa (1), 59 aa (1), 394 aa (2), 396 aa (2), 399 aa (4)
	R3C	10	75.3%-99.8%	70.2%-100% (2)	49 aa (1), 109,aa (1), 396 aa (6), 399 aa (2)
	R4C	11	99.1%-99.8%	89.8%-100% (3)	35 aa (1), 49 aa (1), 302 aa (3), 399 aa (2)
	R1I	10	74.9%-99.7%	69.2%-99.7	106 aa (1), 109 aa (1), 115 aa (1), 269 aa (1), 396 aa (1)
	R4I	11	99.7%-100% (2)	99.8%-100% (9)	282 aa (1)
fliC4 1191 bp 396 aa	R1C	11	75.9%-99.8%	56.3%-99.8%	49 aa (1), 76 aa (1), 237 aa (1), 247 aa (1), 253 aa (1), 273 aa (1), 277 aa (1)
	R4C	10	77.2%-100% (2)	61.7%-100% (2)	59 aa (1), 121 aa (1), 237 aa (1), 244 aa (1), 247 aa (1), 253 aa (1), 276 aa
	R4I	11	77.1%-100% (2)	50.0%-100% (3)	6 aa (1), 237 aa (1), 244 aa (1), 246 aa (1), 247 aa (1), 253 aa (1)
	R6I	10	75.5%-100% (3)	61.5%-100% (5)	247 aa (1)
16S RNA 619 bp	R2C	11	99.7%-100% (7)	NA	NA
	R4C	11	99.4%-100% (8)	NA	NA
	R5C	10	99.7%-100% (6)	NA	NA
	R2I	11	99.5%-100% (7)	NA	NA
	R4I	11	99.4%-100% (6)	NA	NA
	R5I	11	99.4%-100% (4)	NA	NA

NA, not analyzed. Numbers in parentheses represent number of duplicate sequences.

Supplementary Table 4. Statistical analysis of NF-kB signaling data using multi-way analysis of variance (ANOVA) tests*.

A

	PCDNA3	PCDNA3-mTLR5	PCDNA3-rTLR5
Buffer : Sal-FliC 1 μg	1.000	0.863	1.000
Buffer : Sal-FliC 10 μg	1.000	1.000	0.910
Buffer : M5I-FliC3-2 1 μg	1.000	0.943	1.000
Buffer : M5I-FliC3-2 10 μg	1.000	1.000	0.786
Buffer : M5I-FliC3-1 1 μg	1.000	0.896	1.000
Buffer : M5I-FliC3-1 10 μg	1.000	1.000	0.788
Buffer : R4I-FliC3-1 1 μg	0.462	1.000	0.994
Buffer : R4I-FliC3-1 10 μg	0.999	1.000	0.904
Buffer : R1I-FliC3-10 1 μg	1.000	0.993	1.000
Buffer : R1I-FliC3-10 10 μg	1.000	1.000	0.958
Sal-FliC 1 μg : Sal-FliC 10 μg	1.000	1.000	1.000
Sal-FliC 1 μg : M5I-FliC3-2 1 μg	1.000	1.000	1.000
Sal-FliC1 μg : M5I-FliC3-2 10 μg	1.000	1.000	1.000
Sal-FliC 1 μg : M5I-FliC3-1 1 μg	1.000	1.000	1.000
Sal-FliC 1 μg : M5I-FliC3-1 10 μg	1.000	1.000	1.000
Sal-FliC 1 μg : R4I-FliC3-1 1 μg	1.000	1.000	1.000
Sal-FliC 1 μg : R4I-FliC3-1 10 μg	1.000	1.000	1.000
Sal-FliC 1 μg : R1I-FliC3-10 1 μg	1.000	1.000	1.000
Sal-FliC 1 μg : R1I-FliC3-10 10 μg	1.000	1.000	1.000
Sal-FliC 10 μg : M5I-FliC3-2 1 μg	1.000	1.000	1.000
Sal-FliC 10 μg : M5I-FliC3-2 10 μg	1.000	1.000	1.000
Sal-FliC 10 μg : M5I-FliC3-1 1 μg	1.000	1.000	1.000
Sal-FliC 10 μg : M5I-FliC3-1 10 μg	1.000	1.000	1.000
Sal-FliC 10 μg : R4I-FliC3-1 1 μg	1.000	1.000	1.000
Sal-FliC 10 μg : R4I-FliC3-1 10 μg	1.000	1.000	1.000
Sal-FliC 10 μg : R1I-FliC3-10 1 μg	1.000	1.000	1.000
Sal-FliC 10 μg: R1I-FliC3-10 10 μg	1.000	1.000	1.000
M5I-FliC3-2 1 μg : M5I-FliCc3-2 10 μg	1.000	1.000	1.000
M5I-FliC3-2 1 μg : M5I-FliC3-1 1 μg	1.000	1.000	1.000
M5I-FliC3-2 1 μg : M5I-FliC3-1 10 μg	0.995	1.000	1.000
M5I-FliC3-2 1 μg : R4I-FliC3-1 1 μg	1.000	1.000	1.000
M5I-FliC3-2 1 μg : R4I-FliC3-1 10 μg	1.000	1.000	1.000
M5I-FliC3-2 1 μg : R1I-FliC3-10 1 μg	1.000	1.000	1.000
M5I-FliC3-2 1 μg : R1I-FliC3-10 10 μg	1.000	1.000	1.000
M5I-FliC3-2 10 μg : M5I-FliC3-1 1 μg	1.000	1.000	1.000
M5I-FliCc3-2 10 μg : M5I-FliC3-1 10 μg	1.000	1.000	1.000
M5I-FliC3-2 10 μg : R4I-FliC3-1 1 μg	1.000	1.000	0.999

M5I-FliC3-2 10 µg : R4I-FliC3-1 10 µg	1.000	1.000	1.000
M5I-FliC3-2 10 µg : R1I-FliC3-10 1 µg	1.000	1.000	1.000
M5I-FliC3-2 10 µg :R1I-FliC3-10 10 µg	1.000	1.000	1.000
M5I-FliC3-1 1 µg : M5I-FliC3-1 10 µg	0.999	1.000	1.000
M5I-FliC3-1 1 µg : R4I-FliC3-1 1 µg	1.000	1.000	1.000
M5I-FliC3-1 1µg : R4I-FliC3-1 10 µg	1.000	1.000	1.000
M5I-FliC3-1 1 µg : R1I-FliC3-10 1 µg	1.000	1.000	1.000
M5I-FliC3-1 1 µg : R1I-FliC3-10 10 µg	1.000	1.000	1.000
M5I-FliC3-1 10 µg : R4I-FliC3-1 1 µg	1.000	1.000	0.999
M5I-FliC3-1 10 µg : R4I-FliC3-1 10 µg	1.000	1.000	1.000
M5I-FliC3-1 10 µg : R1I-FliC3-10 1 µg	1.000	1.000	1.000
M5I-FliC3-1 10 µg : R1I-FliC3-10 10 µg	1.000	1.000	1.000
R4I-FliC3-1 1 µg : R4I-FliC3-1 10 µg	1.000	1.000	1.000
R4I-FliC3-1 1 µg : R1I-FliC3-10 1 µg	1.000	1.000	1.000
R4I-FliC3-1 1 µg : R1I-FliC3-10 10 µg	1.000	1.000	0.999
R4I-FliC3-1 10 µg : R1I-FliC3-10 1 µg	1.000	1.000	1.000
R4I-FliC3-1 10 µg : R1I-FliC3-10 10 µg	1.000	1.000	1.000
R1I-FliC3-10 1 µg : R1I-FliC3-10 10 µg	1.000	1.000	1.000

B

Buffer	pCDNA3 : pCDNA3-mTLR5	0.740
	pCDNA3 : pCDNA3-rTLR5	0.928
	pCDNA3-mTLR5 : pCDNA3-rTLR5	0.873
Sal-FliC 1 µg	pCDNA3 : pCDNA3-mTLR5	0.199
	pCDNA3 : pCDNA3-rTLR5	0.624
	pCDNA3-mTLR5 : pCDNA3-rTLR5	0.259
Sal-FliC 10 µg	pCDNA3 : pCDNA3-mTLR5	0.590
	pCDNA3 : pCDNA3-rTLR5	0.172
	pCDNA3-mTLR5 : pCDNA3-rTLR5	0.786
M5I-FliC3-2 1 µg	pCDNA3 : pCDNA3-mTLR5	0.248
	pCDNA3 : pCDNA3-rTLR5	0.623
	pCDNA3-mTLR5 : pCDNA3-rTLR5	0.264
M5I-FliC3-2 10 µg	pCDNA3 : pCDNA3-mTLR5	0.499
	pCDNA3 : pCDNA3-rTLR5	0.195
	pCDNA3-mTLR5 : pCDNA3-rTLR5	0.728
M5I-FliC3-1 1 µg	pCDNA3 : pCDNA3-mTLR5	0.063
	pCDNA3 : pCDNA3-rTLR5	0.487
	pCDNA3-mTLR5 : pCDNA3-rTLR5	0.211
M5I-FliC3-1 10 µg	pCDNA3 : pCDNA3-mTLR5	0.503
	pCDNA3 : pCDNA3-rTLR5	0.196
	pCDNA3-mTLR5 : pCDNA3-rTLR5	0.720

R4I-FliC3-1 1 µg	pCDNA3 : pCDNA3-mTLR5	0.816
	pCDNA3 : pCDNA3-rTLR5	0.336
	pCDNA3-mTLR5 : pCDNA3-rTLR5	0.979
R4I-FliC3-1 10 µg	pCDNA3 : pCDNA3-mTLR5	0.496
	pCDNA3 : pCDNA3-rTLR5	0.253
	pCDNA3-mTLR5 : pCDNA3-rTLR5	0.695
R1I-FliC3-10 1 µg	pCDNA3 : pCDNA3-mTLR5	0.319
	pCDNA3 : pCDNA3-rTLR5	0.532
	pCDNA3-mTLR5 : pCDNA3-rTLR5	0.400
R1I-FliC3-10 10 µg	pCDNA3 : pCDNA3-mTLR5	0.484
	pCDNA3 : pCDNA3-rTLR5	0.031
	pCDNA3-mTLR5 : pCDNA3-rTLR5	0.755

*, numbers in this table represent the p value.

Supplementary Table 5. Gene Ontology analysis of the pull-down-enriched proteins.

Category	mFliC3	rFliC3
MF	actin binding	actin binding
	cytoskeletal protein binding	cytoskeletal protein binding
	nucleotide binding	nucleotide binding
	RNA binding	RNA binding
	structural molecule activity	structural molecule activity
	cofactor binding	
	unfolded protein binding	
		exopeptidase activity
CC	actin cytoskeleton	actin cytoskeleton
	cell-cell adherens junction	cell-cell adherens junction
	contractile fiber	contractile fiber
	contractile fiber part	contractile fiber part
	cortical cytoskeleton	cortical cytoskeleton
	cytoskeletal part	cytoskeletal part
	cytoskeleton	cytoskeleton
	intermediate filament	intermediate filament
	intermediate filament cytoskeleton	intermediate filament cytoskeleton
	intracellular non-membrane-bounded organelle	intracellular non-membrane-bounded organelle
	intracellular organelle lumen	intracellular organelle lumen
	melanosome	melanosome
	membrane-enclosed lumen	membrane-enclosed lumen
	mitochondrial envelope	mitochondrial envelope
	mitochondrial inner membrane	mitochondrial inner membrane
	mitochondrial membrane	mitochondrial membrane
	mitochondrial part	mitochondrial part
	mitochondrion	mitochondrion
	myofibril	myofibril
	non-membrane-bounded organelle	non-membrane-bounded organelle
	organelle inner membrane	organelle inner membrane
	organelle lumen	organelle lumen
	organelle membrane	organelle membrane
	pigment granule	pigment granule
	respiratory chain	respiratory chain

	ribonucleoprotein complex	ribonucleoprotein complex
	spliceosome	spliceosome
	adherens junction	
	anchoring junction	
	basolateral plasma membrane	
	endoplasmic reticulum lumen	
	envelope	
	F-actin capping protein complex	
	I band	
	organelle envelope	
	Z disc	
		actin filament bundle
		actomyosin
		cell cortex
		cytosol
		keratin filament
		lysosome
		lytic vacuole
		mitochondrial lumen
		mitochondrial matrix
		soluble fraction
		stress fiber
		vacuole
BP	actin cytoskeleton organization	actin cytoskeleton organization
	actin filament capping	actin filament capping
	actin filament-based process	actin filament-based process
	cytoskeleton organization	cytoskeleton organization
	generation of precursor metabolites and energy	generation of precursor metabolites and energy
	mRNA metabolic process	mRNA metabolic process
	mRNA processing	mRNA processing
	negative regulation of actin filament depolymerization	negative regulation of actin filament depolymerization
	negative regulation of actin filament polymerization	negative regulation of actin filament polymerization
	negative regulation of protein complex assembly	negative regulation of protein complex assembly
	negative regulation of protein polymerization	negative regulation of protein polymerization
	regulation of actin filament depolymerization	regulation of actin filament depolymerization
	RNA processing	RNA processing
	RNA splicing	RNA splicing

electron transport chain	
negative regulation of cellular component organization	
negative regulation of cytoskeleton organization	
negative regulation of organelle organization	
negative regulation of protein complex disassembly	
oxidation reduction	
protein folding	
regulation of actin cytoskeleton organization	
regulation of actin filament length	
regulation of actin filament polymerization	
regulation of actin filament-based process	
regulation of actin polymerization or depolymerization	
regulation of cytoskeleton organization	
regulation of organelle organization	
regulation of protein complex disassembly	
	cellular macromolecular complex subunit organization
	homeostasis of number of cells
	macromolecular complex assembly
	macromolecular complex subunit organization
	protein complex assembly
	protein complex biogenesis
	proteolysis

MF, molecular function; CC, cell component; BP, biology process.

Supplementary Table 6. Number of genes related to actin binding isolated during pull-down that.

GO Terms	mFliC3*	rFliC3*
actin binding	25	25
actin cytoskeleton	16	19
actin cytoskeleton organization	14	18
actin cytoskeleton reorganization	0	3
actin filament binding	0	7
actin filament bundle	0	6
actin filament bundle formation	0	3
actin filament capping	6	6
actin filament organization	0	6
actin filament severing	0	2
actin filament-based process	14	0
F-actin capping protein complex	3	0
negative regulation of actin filament depolymerization	7	6
negative regulation of actin filament polymerization	6	6
regulation of actin cytoskeleton organization	8	6
regulation of actin filament depolymerization	7	6
regulation of actin filament length	8	6
regulation of actin filament polymerization	7	6
regulation of actin filament-based process	8	6
regulation of actin polymerization or depolymerization	8	6

*, number of genes interacted with mFliC3 or rFliC3.

Figure legends

Supplementary Figure 1. Host specificity analysis of SFB 16S rRNA gene sequences.

65 rat SFB 16S rRNA gene sequences were obtained via clone library sequencing. In each sample, duplicate sequences were deleted retaining a single sequence for further analysis. Other human, mouse, chicken and ratSFB 16S rRNA gene sequences were downloaded from NCBI as previously described (1). In total, 74 chicken sequences, 139 human sequences, 112 mouse sequences and 35 rat sequences were used for phylogenetic analysis using MEGA 6 with the neighbor joining algorithm. Green squares represent mouse SFB 16S rRNA gene sequences; red squares represent rat SFB 16S rRNA gene sequences; blue squares represent chicken SFB 16S rRNA gene sequences; purple squares represent human SFB 16S rRNA gene sequences.

Supplementary Figure 2. Host specificity analysis of SFB *fliC1* at nucleotide and deduced amino acid levels.

66 rat SFB *fliC1*, 49 mouse SFB *fliC1* gene sequences were obtained from clone library sequencing. In each sample, duplicate sequences were deleted preserving one sequence for further analysis. Deduced amino acid sequences that were shorter than half of the published SFB FliC genes were removed. In all, 95 *fliC1* nucleotide sequences and 56 deduced FliC1 amino acid sequences were used for phylogenetic analysis. Three mouse SFB *fliC1* sequences and one rat SFB *fliC1* sequence were

downloaded from the NCBI database. Finally, 99 *fliC1* nucleotide sequences (A) and 60 deduced amino acid sequences (B) were aligned and then used to construct a phylogenetic tree using MEGA 6 with the neighbor joining algorithm. Red upward-facing triangles represent rat SFB *fliC1* sequences; the green upward-facing triangles represent mouse SFB *fliC1* sequences.

Supplementary Figure 3. Staining of *E. coli* BL21 and *Salmonella* CVCC519 with DAPI and SFB FliC3 antibody.

The bacterium *E. coli* BL21 and *Salmonella* CVCC519 were fixed in 4% paraformaldehyde solution and sequentially hybridizedwith a rabbit anti-FliC3 primary antibody and DAPI. An Olympus IX83 was used to observe the results. A and D represent the DAPI stain results of *E. coli* BL21 and *Salmonella* CVCC519, respectively. B and E represent the SFB FliC3 antibody stain results of *E. coli* BL21 and *Salmonella* CVCC519, respectively. C and F are merged images.

Supplementary Figure 4. Expression of SFB FliC3 in mouse ileum mucosa and cecal content using FISH and IHC analysis.

Ileal mucosa and cecal contents were fixed in 4% paraformaldehyde solution, and sequential hybridization with a SFB-specific oligonucleotide probe, rabbit anti-FliC3 primary antibody and FITC goat anti-rabbit IgG (H+L) antibody, DAPI was then used to stain the nuclei. A Leica fluorescent microscopeDM2500 was used to observe the results. A and E, represent the DAPI stain result. B and F, represent the

results of hybridization with SFB-specific oligonucleotide probe. C and G, represent the results of hybridization with rabbit anti-FliC3 primary antibody and FITC goat anti-rabbit IgG (H+L) antibody. D and H, represent the overlapped signal between B and C. Arrow represents the presence of SFB.

Supplementary Figure 5. Relative location of SFB in gut epithelial tissue.

Ileal (A-D) and colon (E-H) tissue samples were fixed in 4% paraformaldehyde and prepared as frozen sections. After sequential hybridization with primary antibody (mouse anti-muc2 and rabbit anti-mFliC3 antibody) and the secondary antibody (donkey anti-mouse IgG (H+L) and FITC goat anti-rabbit IgG (H+L) antibody), DAPI was then used to stain nuclei. A two photon confocal microscope (ZEISS LSM 800) were used to observe the cells. A and E show the DAPI stain results. B and F show the hybridization results with rabbit anti-mFliC3 primary antibody. C and G show the results of hybridization with mouse anti-muc2 primary antibody. D represents the overlapped signal between B and C. H represents the overlapped signal between F and G.

Supplementary Figure 6. Identification of genes related to flagellar assembly using LC/MC/MC in mouse gut samples.

Ileum mucosal and cecal contents protein on the SDS-PAGE gel in the 37-50 kD region were cut from the gel and sent for LC/MC/MC identity. Proteins relating to flagellar assembly among identified SFB proteins are marked with a red box.

Supplementary Figure 7. Peptides identified from purified mFliC3 bands.

Bands 1 and 2 (as illustrated in Figure 5) were cut from purified mFliC3 protein. Peptides identified from these two bands were then mapped FliC3 amino acid sequences of SFB-mouse-Japan (BioProject: PRJDA66727).

Supplementary Figure 8. Phylogenetic analysis of the TLR5 genes at the nucleotide and deduced amino acid levels.

50 chicken TLR5, 6 mouse TLR5 and 7 rat TLR5 nucleotide sequences were downloaded from the NCBI database. Nucleotide sequences (A) and deduced amino acid sequences (B) were aligned and then used to construct a phylogenetic tree using MEGA 6 with the neighbor joining algorithm. Green squares represent mouse TLR5 sequences; the red squares represent rat TLR5 sequences; the blue squares represent chicken TLR5 sequences.

Supplementary Figure 9. RFliC3 enriched proteins that related to the lysosome pathway.

Mouse mucosal proteins which were collected by pull-down with SFB FliC3 were identified using LC/MS/MS and sent then for KEGG pathway analysis.rFliC3 enriched proteins that related to the lysosome pathway were marked with red boxes.

Supplementary Figure 10. Degradation of mFliC3 and rFliC3 by rat ileum tissue proteins.

Purified mFliC3 and rFliC3 were separately allowed to interact *in vitro* with rat mucosal proteins. (A) samples of these interaction mixtures were collected at different times and assessed using SDS-PAGE and immunoblotting with an anti-His antibody. (B) The concentrations of the protein bands were quantified by Image J (B).

Supplementary Figure 11. Visualization of mutant variable regions of the deduced SFB FliC3 and FliC4 amino acid sequences.

52 rat SFB *fliC3*, 42 rat SFB *fliC4*, 66 mouse SFB *fliC3* and 51 mouse SFB *fliC4* gene sequences were obtained from clone library sequencing. In each sample, duplicate sequences were deleted except one with one sequence preserved for further analysis. Deduced amino acid sequences which were shorter than half of the published SFB FliC genes were removed. In all, 130 SFB FliC3 and FliC4 sequences obtained in this study and 8 SFB FliC3 and FliC4 sequences downloaded from the NCBI database were used for alignment. JProfileGrid software was used for multiple sequence alignment visualization. Different colors represent sequence homology at the site for each amino acid. The number represents the percentage chance for each amino acid.

Reference:

1. Yin Y, Wang Y, Zhu L, Liu W, Liao N, Jiang M, Zhu B, Yu HD, Xiang C, Wang X. 2013. Comparative analysis of the distribution of segmented filamentous bacteria in humans, mice and chickens. ISME J 7:615-21.