IAI Journals.ASM.org

Evolutionary Adaptation of an AraC-Like Regulatory Protein in *Citrobacter rodentium* and *Escherichia* Species

Aimee Tan,[a]* Nicola K. Petty,[b] Dianna Hocking,[a] Vicki Bennett-Wood,[a] Matthew Wakefield,[c] Judyta Praszkier,[a] Marija Tauschek,[a] Ji Yang,[a] Roy Robins-Browne[a,d]

Department of Microbiology and Immunology, University of Melbourne, Peter Doherty Institute for Infection and Immunity, Victoria, Australia[a]; ithree Institute, University of Technology, Sydney, NSW, Australia[b]; Walter Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia[c], Murdoch Children's Research Institute, Royal Children's Hospital, Parkville, Victoria, Australia[d]

The evolution of pathogenic bacteria is a multifaceted and complex process, which is strongly influenced by the horizontal acquisition of genetic elements and their subsequent expression in their new hosts. A well-studied example is the RegA regulon of the enteric pathogen *Citrobacter rodentium*. The RegA regulatory protein is a member of the AraC/XylS superfamily, which coordinates the expression of a gene repertoire that is necessary for full pathogenicity of this murine pathogen. Upon stimulation by an exogenous, gut-associated signal, namely, bicarbonate ions, RegA activates the expression of a series of genes, including virulence factors, such as autotransporters, fimbriae, a dispersin-like protein, and the *grlRA* operon on the locus of enterocyte effacement pathogenicity island. Interestingly, the genes encoding RegA homologues are distributed across the genus *Escherichia*, encompassing pathogenic and nonpathogenic subtypes. In this study, we carried out a series of bioinformatic, transcriptional, and functional analyses of the RegA regulons of these bacteria. Our results demonstrated that *regA* has been horizontally transferred to *Escherichia* spp. and *C. rodentium*. Comparative studies of two RegA homologues, namely, those from *C. rodentium* and *E. coli* SMS-3-5, a multiresistant environmental strain of *E. coli*, showed that the two regulators acted similarly *in vitro* but differed in terms of their abilities to activate the virulence of *C. rodentium in vivo*, which evidently was due to their differential activation of *grlRA*. Our data indicate that RegA from *C. rodentium* has strain-specific adaptations that facilitate infection of its murine host. These findings shed new light on the development of virulence by *C. rodentium* and on the evolution of virulence-regulatory genes of bacterial pathogens in general.

Increases in the amount and availability of DNA sequence data have provided valuable insights into the ongoing speciation and evolution of bacteria. In particular, these data have revealed the frequency and ubiquity of horizontal gene transfer and how these events may have contributed to the development of bacterial pathogens (1–11). For example, the acquisition of the pPla and pMT1 plasmids catalyzed speciation of *Yersinia pestis* from *Yersinia pseudotuberculosis* (10), and the recent emergence of a new pathogenic strain of *Escherichia coli*, which caused a dramatic foodborne disease outbreak in Germany in 2011, is attributed to the horizontal transfer of a number of virulence determinants, including Shiga toxin, various adhesins and autotransporters, and the pESBL C227-11 plasmid (12–14).

However, the acquisition of virulence determinants alone is often not sufficient to make a bacterium pathogenic. For example, the murine pathogen *Citrobacter rodentium* contains a large pathogenicity island known as the locus of enterocyte effacement (LEE PAI), which is required for colonic hyperplasia and the formation of attaching and effacing (A/E) lesions. Evidence indicates that the island has been horizontally mobilized into *C. rodentium* (15) and different lineages of enteropathogenic (EPEC) and enterohemorrhagic (EHEC) *E. coli* (16–18), as well as into *Escherichia albertii*, *Shigella boydii*, and *Salmonella enterica* subsp. *salamae* (19, 20). However, transfer of the island does not always recapitulate the A/E phenotype (21, 22). Morevoer, the LEE PAI and the phage-borne Shiga toxin (*stx*) genes have also been found in *E. coli* isolates from healthy ruminants (23–25) and birds (26). A possible explanation for this lack of pathogenicity may be dysfunction in gene regulation in new hosts and genetic backgrounds, leading to the inability to stably or intimately colonize the host in which disease occurs or the inability to activate the virulence determinants at the appropriate time. The LEE PAI is internally regulated by the Ler, GrlA, and GrlR proteins (27) but generally also requires a LEE-external regulator to initiate transcription of its genes (28).

In *C. rodentium*, the LEE PAI is externally regulated by the RegA protein via its effects on *grlA* (29). RegA is a member of the AraC/XylS family of proteins and shares similarity with a subset of virulence regulators, including PerA from enteropathogenic *E. coli* (30–32), Rns from enterotoxigenic *E. coli* (33–36), AggR from enteroaggregative *E. coli* (37–40), ToxT from *Vibrio cholerae* (41–44), and VirF from *Shigella flexneri* (45–47). These proteins typi-

Received 27 September 2014 **Returned for modification** 22 October 2014
Accepted 16 January 2015

Accepted manuscript posted online 26 January 2015

Citation Tan A, Petty NK, Hocking D, Bennett-Wood V, Wakefield M, Praszkier J, Tauschek M, Yang J, Robins-Browne R. 2015. Evolutionary adaptation of an AraC-like regulatory protein in *Citrobacter rodentium* and *Escherichia* species. Infect Immun 83:1384–1395. doi:10.1128/IAI.02697-14.

Editor: A. J. Bäumler

Address correspondence to Roy Robins-Browne, r.browne@unimelb.edu.au.

* Present address: Aimee Tan, Institute for Glycomics, Griffith University, Gold Coast, QLD, Australia.

M.T., J.Y., and R.R.-B. contributed equally to this work.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/IAI.02697-14.

doi:10.1128/IAI.02697-14

TABLE 1 Bacterial strains and plasmids used in this study

Bacterial strain or plasmid	Comment or description[a]	Source or reference
Strains		
ICC169	Nal^r derivative of *C. rodentium* ICC168	58
EMHI	*C. rodentium regA* deletion strain (ICC169 Δ*regA*::Kan^r)	52
MC4100	*E. coli* K-12 strain [Δ(*argF–lac*)*U169 rpsL150 relA araD139 flb-5301 deoC1 ptsF25*]	59
SMS-3-5	Environmental *E. coli* strain	60
Plasmids		
pACYC184	Low-copy-number cloning vector; Tc^r Cm^r	New England Biolabs
pAT6	ICC169 RegA complementation plasmid, pACYC184 backbone	This study
pAT9	SMS-3-5 RegA complementation plasmid, pACYC184 backbone	This study
grlRA-1	ICC169 *grlRA* promoter transcriptional fusion construct, pMU2385 backbone	29
kfcC-lacZ	ICC169 *kfcC* promoter transcriptional fusion construct, pMU2385 backbone	54
kfcC-lacZ mutant	ICC169 *kfcC* promoter proximal GATATAA motif deletion transcriptional fusion construct, pMU2385 backbone	50
regB-lacZ	ICC169 *regB* promoter transcriptional fusion construct, pMU2385 backbone	50
regB-lacZ mutant	ICC169 *regB* promoter GATATAA motif deletion transcriptional fusion construct, pMU2385 backbone	50
pMU2385	Promoterless *lacZ* transcriptional fusion plasmid	61
pMU2386	Promoterless *lacZ* translational fusion plasmid	62
pMUAT6	SMS-3-5 *regA* translational fusion construct, pMU2386 backbone	This study
pMUAT7	ICC169 *regA* translational fusion construct, pMU2386 backbone	This study

[a] Nal^r, nalidixic acid resistance; Kan^r, kanamycin resistance; Tc^r, tetracycline resistance; Cm^r, chloramphenicol resistance.

cally bind AT-rich consensus binding sites (30, 36, 48–50) and, consequently, can relieve H-NS repression to activate gene transcription (44, 47, 50, 51).

In *C. rodentium*, RegA is a key virulence regulator necessary for colonization and full pathogenicity for mice, its natural host (52, 53). In addition to its role in activation of the *grl* regulator (29), RegA regulates a global gene repertoire, including virulence genes, such as *adcA* and the *kfc* operon, which encode an autotransporter and an adhesin, respectively, and the *aap*-encoded dispersin homologue and associated *aat* transport system (52, 54). RegA regulation is enhanced by bicarbonate ions (50, 54), which act as a cofactor for the protein and an environmental stimulus that signals transition of the bacterium into the small intestine (55, 56).

Interestingly, however, homologues of *regA* have recently been identified in both pathogenic and nonpathogenic strains of bacteria. Additionally, recent studies on the RegA homologue RegR, from rabbit-enteropathogenic *E. coli* strain E22, demonstrated differences in regulatory targets (57). Together, these data suggest that the role of RegA as a key virulence and LEE PAI regulator may be an evolutionary adaptation specific to *C. rodentium*. The identification of various RegA-like regulatory systems provides an opportunity to examine how virulence regulation might have evolved in enteric bacteria.

To examine this, we performed a comparative analysis of all the RegA homologues identified in publicly available genome sequences and analyzed the comparative functions of the RegA homologues of *C. rodentium* ICC169 and *E. coli* strain SMS-3-5. Collectively, these studies indicated that RegA is a horizontally transferred transcriptional regulator that has been coopted to regulate global virulence genes in *C. rodentium*.

MATERIALS AND METHODS

Bacterial strains, plasmids, media, and primers. The bacterial strains and plasmids used in this study are listed in Table 1. Unless stated otherwise, bacteria were grown at 37°C in Luria-Bertani broth (LB) or on Luria agar (LA) plates supplemented with the appropriate antibiotics. The PCR primers used in this study are listed in Table 2.

Bioinformatic and phylogenetic analyses. RegA homologues were identified by BLASTP and tBLASTn analyses using the ICC168 protein sequence (accession no. CBG90830.1) to query the NCBI nonredundant protein database. Protein sequence comparisons were performed using the MATGAT program (63). In the case of homologues with truncated lengths, identity and similarity were calculated based on the sequence available with no gap penalties. Multilocus sequence typing (MLST) analysis was performed on these strains according to the guidelines for the st7 scheme at EcMLST (http://www.shigatox.net/ecmlst/protocols/mlst_primers.html [shigatox.net]), with sequence data obtained by BLASTn searches of the respective genome sequences in the NCBI databases. Where necessary, this was supplemented by the sequencing of allele amplicons from the relevant strains using the method described at the shigatox.net website. Sequence data from *S. enterica* LT2 and various strains of *Citrobacter* spp. were also obtained to contextualize the distance between *C. rodentium* and the various *Escherichia* sp. strains containing *regA*. Additional data for this assay were from fully annotated *E. coli* and *Shigella* strains lacking *regA* reported previously (64). Phylogenetic trees were gen-

TABLE 2 PCR primers used in this study

Name	Sequence (5′–3′) or source[a]
AT120	AA**GGATCC**TACAGAGGTAAGGGATAACTCATTG
AT122	AA**GTCGAC**GCAACGCTATTTAATCCTCCGG
AT139	AT**GGATCC**CTCACGTTGAATTTTCGAGTTTACAC
AT140	TT**GTCGAC**GAATTCGCTGTTCAGGCAGC
AT141	**AAGCTT**CATGGTAAGAGTGTAAGAGAGATTTC
AT142	TT**GGATCC**TTTGTACAAGTAACTCGACTATTCAT
AT149	A**GGATCC**GCAGCCATTGCAGTAGAAGTC
AT150	T**AAGCTT**CAGATATCAACCGCATCCCG
AT151	**AAGCTT**GGATTCAGAATTAAGGCGATAGCG
AT152	**GGATCC**TTACTTAAAGCACTTTGACGATTCAT
MLST primers	http://www.shigatox.net/ecmlst/protocols/mlst_primers.html

[a] Boldface type indicates engineered restriction enzyme sites.

erated by the MEGA program (65), using the neighbor-joining algorithm. Clustal analyses were run using the ClustalW2 multiple-sequence alignment tool at EBI (http://www.ebi.ac.uk/Tools/msa/clustalw2/). Local genome comparisons were carried out using BLAST in the Easyfig program (66). Where necessary, genome sequence manipulation and annotation were carried out by using the Artemis program (67).

Construction of *trans*-complementing plasmids pAT6 and pAT9. For *trans*-complementation of bacterial strains, 1,157- and 995-bp DNA fragments containing *regA* were amplified from *C. rodentium* ICC169 and *E. coli* SMS-3-5 genomic DNA using primer pairs AT120/AT122 and AT139/AT140, respectively (Table 2). These fragments were each cloned into pCR2.1-TOPO and confirmed by sequencing. The fragments were then excised by digestion with BamHI/SalI and ligated into BamHI/SalI-digested pACYC184 to generate pAT6 (for $RegA_{CR}$) and pAT9 (for $RegA_{SMS}$).

Construction of *lacZ* fusion plasmids. A 393-bp fragment consisting of the upstream region and the first nine triplets of the coding sequences of $regA_{SMS}$ were amplified from *E. coli* SMS-3-5 genomic DNA by using primer pair AT141/AT142. The fragment was cloned into pCR2.1-TOPO and confirmed by sequencing. It was then excised by digestion with HindIII/BamHI and inserted into the single-copy-number translational-fusion vector pMU2386 to generate pMUAT6. The above-described plasmid construction method was also used to generate the translational-fusion vector pMUAT7 with primer pair AT151/AT152 and *C. rodentium* ICC169 genomic DNA. The transcriptional fusions used in this study were sourced from other studies as follows: grlRA-1 (29); *kfc-lacZ* (54); and *kfc-lacZ* mutant, *regB-lacZ*, and *regB-lacZ* mutant (50).

β-Galactosidase assays. β-Galactosidase assays were carried out as described by Miller (68). Unless otherwise indicated, cells were grown in LB, in the presence of 45 mM sodium bicarbonate, to mid-log phase (optical density at 600 nm [OD_{600}] = 0.5 to 0.8). The results shown are from at least three independent assays.

Infection of mice. To assess the abilities of *C. rodentium* strains to colonize the mouse intestine, 4-week-old male C57BL/6 mice were infected by oral gavage with approximately 1×10^9 CFU of a test or control strain. Fecal samples were collected daily, diluted in PBS, and plated onto LA plates containing nalidixic acid (50 μg/ml) and chloramphenicol (10 μg/ml) for assessment of intestinal colonization over 15 days. At the peak of infection (day 10), mice from each test and control group were euthanized by CO_2 inhalation, and colon tissue (anus to cecum) was taken, weighed, and bisected for enumeration of bacteria and histological analysis. Colon sections were stained with hematoxylin and eosin, examined using a Leica SMI4000B inverted microscope at ×100 magnification, and measured with the micrometer in the Leica application suite software. To compare the virulence of *C. rodentium* strains, 3-week-old male C57BL/6 mice were infected by oral gavage with approximately 1×10^9 CFU of a test or control strain. Mice were culled if sufficient signs of physiological stress or illness were observed, including severe weight loss, dehydration, rectal prolapse, and/or postural or behavioral defects. All experimental procedures using animals were approved by the University of Melbourne Animal Experimentation and Ethics Committee and performed in accordance with the guidelines for animal experimentation of the Australian National Health and Medical Research Council.

Antisense *C. rodentium* ICC169 microarray. An antisense oligonucleotide microarray was designed by using the eArray platform (Agilent Technologies). The arrays contained open reading frames (ORFs) representing all gene predictions for *C. rodentium* strain ICC168 (accession number FN543502) (58). Each ORF was represented by at least three different oligonucleotides.

RNA isolation and labeling. RNA was isolated from bacterial cells grown in LB supplemented with chloramphenicol (10 μg/ml) and with 45 mM sodium bicarbonate to an OD_{600} of 0.8 to 0.9. One volume of cells was incubated with 2 volumes of RNAprotect solution (Qiagen) and processed according to the manufacturer's instructions. Cell lysis and RNA preparation were carried out using the FastRNA Pro Blue kit (Qbiogene Inc.). After treatment with RNase-free DNase I (Qiagen), the RNA was further purified using the RNeasy MinElute kit (Qiagen). A total of 5 μg of RNA was labeled either with Cy-5-ULS or with Cy-3-ULS as described in the Kreatech ULS labeling manual (Kreatech Diagnostics), followed by determination of the RNA quality and degree of labeling with an Agilent 2100 Bioanalyzer and an ND-1000 spectrophotometer (NanoDrop Technologies).

Fragmentation, microarray hybridization, scanning, and analysis. Fragmentation, hybridization, and scanning of the microarray were performed at the Australian Genome Research Facility Ltd. (AGRF) (Melbourne, Australia), as described previously (52). Normalization and data analysis were performed using the limma package in Bioconductor (69–71). Genes were considered differentially expressed if they showed an average change of >2-fold with an adjusted *P* value of <0.05.

Statistical methods. Graphs and statistical analyses were done using GraphPad Prism 5.0 and Instat software (GraphPad Software). Data were analyzed by using one-way analysis of variance (ANOVA) with the Tukey-Kramer posttest or by unpaired Student's *t* test, two tailed. *P* values of <0.05 were taken to indicate statistical significance.

Microarray data accession number. The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (GEO) (72) and are accessible through GEO series accession number GSE57689.

RESULTS AND DISCUSSION

Identification and distribution of RegA homologues. RegA homologues were identified by using BLASTP and tBLASTn searches of the NCBI and Broad Institute databases with the *C. rodentium* RegA protein sequence (here referred to as $RegA_{CR}$) and were deemed homologous if they showed a minimum of 60% amino acid identity to $RegA_{CR}$. From these searches, we identified 19 putative RegA homologues, including the RegR protein from enteropathogenic *E. coli* strain E22 (57). An additional RegA homologue was identified in the *Escherichia* clade IV strain H605 (David Gordon, personal communication) (see Table S1 in the supplemental material). A minimum of 62.2% identity was seen between all these RegA homologues, with an overall average identity to the $RegA_{CR}$ homologue of 69.5%. In comparison, the next most similar protein was from *E. coli* NC101 (GenBank accession no. gb|EFM54601.1), which shared only 46.4% identity with $RegA_{CR}$ and a maximum of 49.6% identity with any other RegA homologue (see Table S2 in the supplemental material). No *regA* pseudogenes were identified in searches, suggesting that RegA may be under positive selection.

To examine the distribution of the homologues, the respective host strains were analyzed by MLST using a 7-allele scheme previously used to categorize atypical EPEC (73) and ETEC (64) strains. Data from this analysis (Fig. 1A) show the basic phylogeny of the *Enterobacteriaceae*, with discrete clustering of strains of the *Escherichia* clade separate from *E. coli* and *Shigella*, and the polyphyletic nature of the genus *Citrobacter*, as previously noted (58, 74). These data also demonstrate the sporadic distribution of the *regA* gene among *E. coli* strains and limited clustering of *regA*-containing strains in *Escherichia* clades III, IV, and V. However, the clustering we observed in the non-*E. coli Escherichia* clades may be biased due to the limited number of genomes sequenced from each clade. The low frequency and wide distribution of *regA* across the MLST phylogenetic tree suggest that *regA* is likely to have been transferred horizontally.

Phylogeny of *regA* and RegA. Analysis of *regA* phylogeny revealed that two major groups of *regA* homologues exist: the first group contains homologues from the *Escherichia* strains SE15, AA86, M605, EC958, H605, TW09276, W26, E22, PCN033, and 97.0259 (here called group 1), which share a minimum of 93.5%

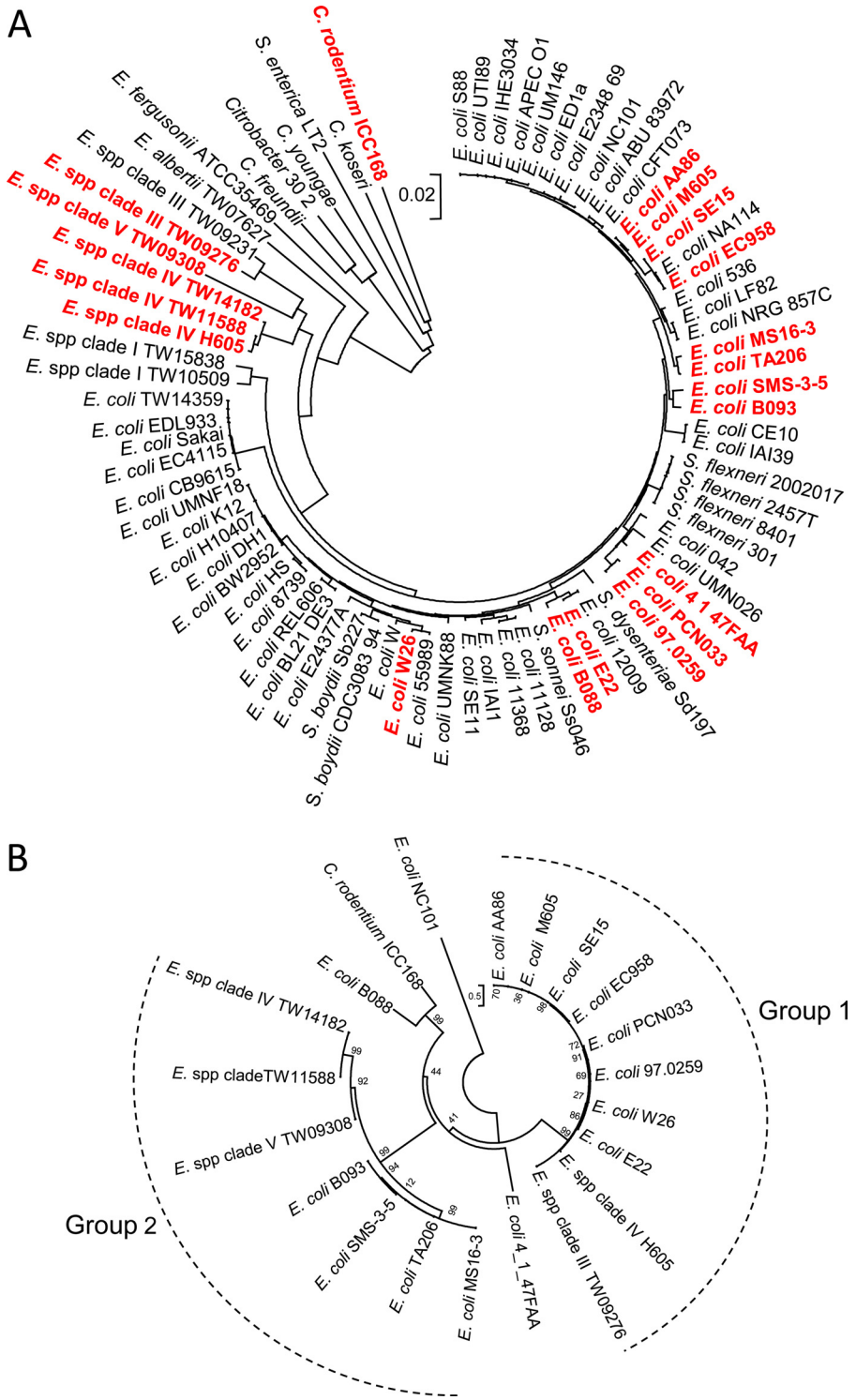

FIG 1 Phylogenetic analyses of *Enterobacteriaceae* and *regA*. (A) Phylogenetic analysis of concatenated multilocus sequence alleles of *Escherichia* (*E.*), *Citrobacter* (*C.*), and *Shigella* (*S.*) strains, using the 7-allele scheme described previously (64, 73). Strains containing *regA* are depicted in red. The scale bar represents 0.02 nucleotide changes per position. (B) Phylogenetic analysis of *regA* nucleotide sequences and NC101 *regA*-like sequence. Group 1 and group 2 *regA* homologues are indicated by dashed lines. The scale bar represents 0.5 nucleotide changes per position. The phylogenetic trees were generated in the MEGA5 program using the neighbor-joining method with 1,000 bootstrap replications. Bootstrap values of less than 100% are shown at the internal nodes.

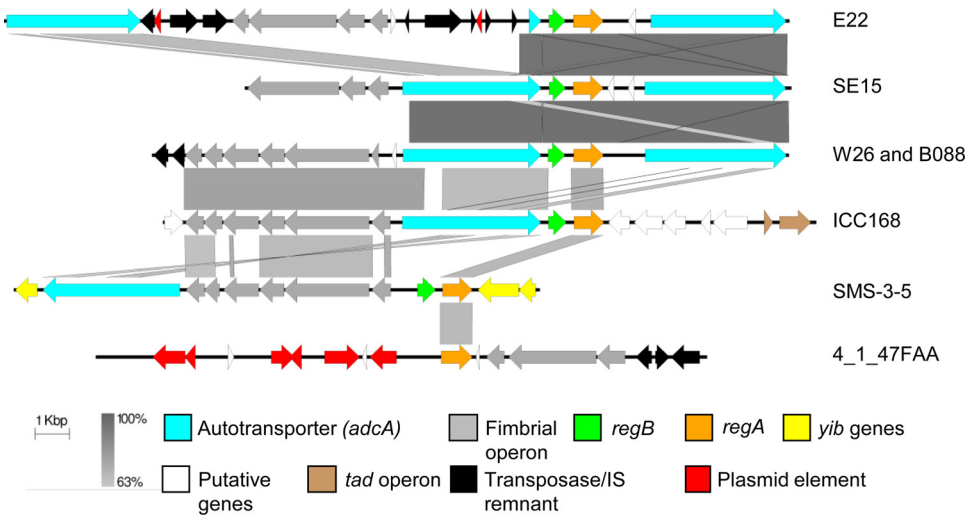

FIG 2 Pairwise comparison of representative *regA* genetic loci. The *regA* genetic loci from *E. coli* strains E22, SE15 (group 1 representative), W26, and B088; *C. rodentium* ICC168; and *E. coli* strains SMS-3-5 (group 2 representative) and 4_1_47FAA were compared using Easyfig 2.0 software (66). Coding regions are represented by colored arrows, with the predicted functions shown at the bottom. The levels of nucleotide identity (BLASTn) between adjacent strains are indicated by the gray shading between the loci and correspond to the scale bar (bottom left).

identity (see Table S2 in the supplemental material), and the other group contains homologues from the *Escherichia* strains TW14182, TW11588, TW09308, TA206, B093, SMS-3-5, and MS16-3 (here called group 2), which share a minimum of 83.7% identity (see Table S2 in the supplemental material). Three homologues, from *E. coli* B088, *E. coli* 4_1_47FAA, and *C. rodentium* ICC168, did not cluster with either group. Of these, the *E. coli* B088 and *C. rodentium* ICC168 homologues clustered together (83.3% identity) (see Table S2 in the supplemental material).

Comparison of these data with the MLST analyses indicated that the evolutionary history of *regA* was only partially consistent with the evolutionary histories of the host strains. Congruous *regA* gene phylogeny and MLST phylogeny was seen with the *E. coli* strains SE15, EC958, AA86, and M605 and with the strains MS16-3, TA206, SMS-3-5, and B093. Elsewhere, however, we observed highly similar *regA* homologues in more evolutionarily distinct strains, e.g., with *E. coli* B088 and *C. rodentium* ICC168, and also with the *Escherichia* clade III strain TW09276 and *E. coli* SE15 (Fig. 1B). These findings support the notion that *regA* was horizontally mobilized in its evolutionary history.

Further evidence of horizontal transfer of the *regA* gene was revealed by comparative analyses of the *regA* locus. In general, we found that most strains with similar *regA* homologues have similar *regA* loci that were consistent with the division of these strains into groups 1 and 2. For further analysis we used the SE15 and SMS-3-5 strain genomes as representatives of group 1 and 2 loci, respectively, as they had the most complete genome sequences available. Full group comparisons are shown in Fig. S1 and S2 in the supplemental material.

The *regA* locus in most strains containing group 1 *regA* homologues comprised a core unit consisting of the *adcA* adhesin/autotransporter gene (52), the uncharacterized *regB* gene (50), the *regA* gene, and a second putative autotransporter gene. This core unit was also typically transcribed divergently from a gene cluster with fimbrial domains. In SE15 and most group 1-containing strains, the cluster comprised three genes, with *fimA*, *fimC*, and *fimD* domains, respectively, and the entire locus was located between the *dcuS* and *lysU* genes (Fig. 2, SE15). However, fimbrial and locus variations were seen in the E22, W26, and B088 strains. In E22, the core unit was rearranged, with the *adcA* gene located upstream of the *sef* fimbrial genes (57), and numerous transposable elements and remnants were present (Fig. 2, E22). The locus also appeared to be in a different genetic location, suggesting that localized genome reorganization or horizontal transfer may have occurred.

In *E. coli* W26 and B088, the three-gene fimbrial cluster was replaced by a fimbrial operon analogous to the *C. rodentium kfc* operon. Interestingly, the W26 and B088 strains had identical organizations and high genetic identity in this locus, with the exception of the *regB-regA* sequence, suggesting that horizontal replacement of the *regA* gene may have occurred (see Fig. S1B in the supplemental material). This core locus organization of *kfc-adcA-regB-regA* was also shared with *C. rodentium* (Fig. 2). However, *C. rodentium* ICC168 lacks a second autotransporter gene and is situated upstream of an operon (ROD_41381 to ROD_41451) encoding Flp pili involved in colonization and biofilm formation (Fig. 2, ICC168). Elsewhere, this operon is found on a widespread mobile genomic island in disparate genera, including *Aggregatibacter*, *Haemophilus*, *Pseudomonas*, and *Yersinia* (75), suggesting that the region may facilitate integration of foreign DNA.

In SMS-3-5 and other strains with group 2 homologues, the *regA* locus comprised the *regBA* genes transcribed divergently from a fimbrial operon homologous to the *kfc* operon and an *adcA* autotransporter homologue (Fig. 2, SMS-3-5). Interestingly, these loci were located between the putative *yibF* and *yibHI* genes, in place of the *yibAJGV* locus and an *rhsA* element found in this location in *E. coli* K-12. Rhs elements are known to be involved in gene duplication and genome rearrangement events (76, 77), suggesting that the region may be genetically mobile.

The *regA* locus in *E. coli* 4_1_47FAA bore no genomic synteny to any other *regA* locus and lacked any *regB* and *adcA* homologues (Fig. 2, 4_1_47FAA). The *regA* gene in this strain is flanked by numerous transposable-element genes and a genetic region homologous to the *E. coli* F plasmid, with genes encoding the RepFIB replication protein A and site-specific recombinase located imme-

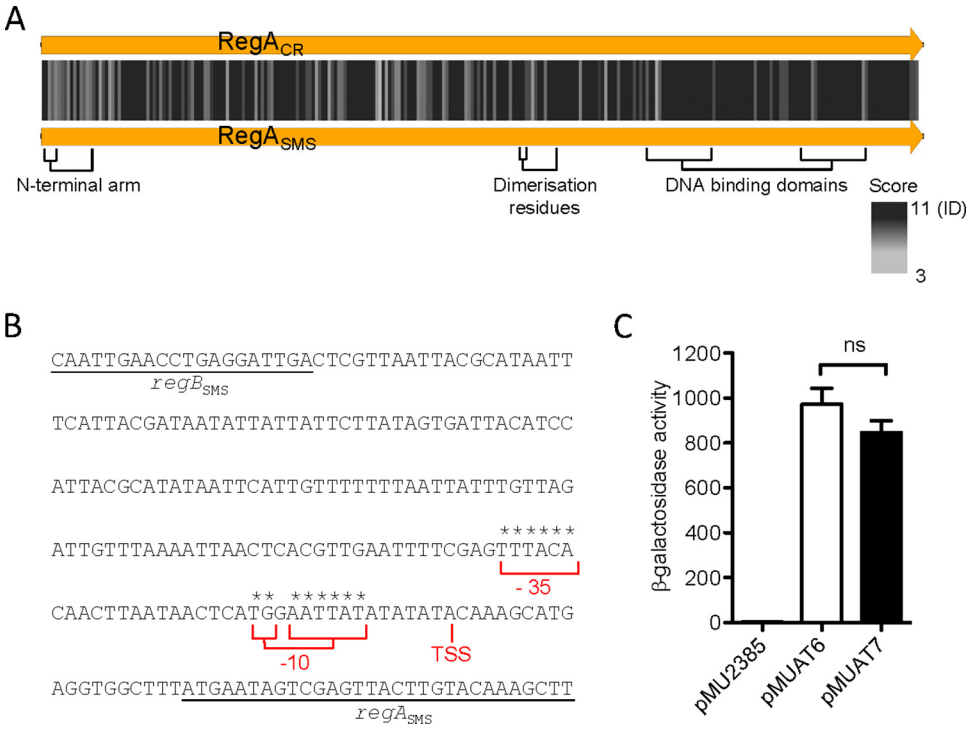

FIG 3 Analysis of RegA from *E. coli* SMS-3-5 (RegA$_{SMS}$). (A) Comparison of RegA$_{CR}$ (top arrow) and RegA$_{SMS}$ sequences, with locations of residues important for the N-terminal-arm function, dimerization, and DNA binding domains indicated. Sequence similarity is represented by the lines of gray shading between proteins, according to the conservation score scale (bottom right; generated in Jalview [91]), with dark-gray lines representing amino acid identity (ID) (a score of 11) and light-gray lines representing relative amino acid mismatch between the two proteins (a score of 3). (B) *regA*$_{SMS}$ promoter and regulatory sequences. The coding regions of the 3′ end of the respective *regB* gene and the 5′ end of *regA* are underlined. The locations of sequences corresponding to the experimentally determined *regA*$_{CR}$ promoter (−35 and −10 box with extended TG motif) (50) and the transcriptional start site are indicated in red, with conserved residues indicated by asterisks. (C) Demonstration of the initiation of transcription and translation of *regA*$_{SMS}$ and *regA*$_{CR}$ fused to a promoterless *lacZ* gene. The data shown are β-galactosidase activities (Miller units) in *C. rodentium* ICC1699, grown to mid-log phase in LB, and are the means and SD of three independent experiments. ns, $P > 0.05$; Student's *t* test, two tailed.

diately upstream of *regA* (accession no. NC_002483.1) (Fig. 2, 4_1_47FAA; see Fig. S3 in the supplemental material, F plasmid comparison). These genes and the lack of synteny seen with other *regA* loci suggest that *regA* has been mobilized into this location.

Collectively, these analyses indicated that limited levels of vertical inheritance of *regA* may have occurred among strains within groups 1 and 2, where conserved genomic synteny and consistent MLST and *regA* phylogeny were observed. On the other hand, marked inconsistencies between MLST, *regA* phylogenies, and *regA* locus organization suggested that *regA* has been horizontally transmitted multiple times in its evolutionary history. The relatively few extant homologues suggested that transfer of the *regA* locus is an infrequent event. In addition, the distribution and number of *regA* homologues in *Escherichia* strains suggested *regA* may have originated in the genus (although this analysis may be limited by the relatively few non-*E. coli* genomes available). If so, the *regA* locus may represent yet another genetic element that has been horizontally transferred into *C. rodentium* from *E. coli* (78).

Functional analysis of RegA$_{CR}$ and RegA$_{SMS}$. The finding of RegA homologues in pathogenic and nonpathogenic *Escherichia* strains (see Table S1 in the supplemental material) suggested that the well-studied virulence association of RegA$_{CR}$ in *C. rodentium* may not be evolutionarily conserved and may have varied in different genetic backgrounds. To study how the *C. rodentium* background may have affected the evolution of RegA$_{CR}$, the RegA$_{CR}$ homologue from *C. rodentium* ICC169 (a nalidixic acid-resistant derivative of ICC168) was compared to the RegA homologue from the environmental *E. coli* strain SMS-3-5 (here referred to as RegA$_{SMS}$). These proteins share 66.3% identity and 82.6% amino acid similarity, with much of the divergence in their N-terminal domains, in regions responsible for ligand binding (55) (Fig. 3A). Some differences were also apparent in the C-terminal DNA binding domains, but these motifs were mostly conserved. In addition, the specific residues important for the dimerization of the RegA protein were conserved. Examination of the *regA*$_{SMS}$ promoter region showed that it shares identical promoter sequences with *regA*$_{CR}$ (experimentally determined [50]) (Fig. 3B), indicating a high likelihood that RegA$_{SMS}$ is produced. As a first step to confirm transcription and translation of RegA$_{SMS}$, this promoter region and the first nine codons of *regA*$_{SMS}$ were amplified and cloned into pMU2386, the translational *lacZ* fusion plasmid, to produce pMUAT6. This construct was transformed into *C. rodentium* ICC169, which was then grown to mid-log phase and assayed for β-galactosidase activity. The assays demonstrated that pMUAT6 had 972 ± 72 Miller units (U) of enzyme activity, compared to 2.7 ± 0.2 U for the empty vector (pMU2385), confirming transcription and translation initiation of *regA*$_{SMS}$ (Fig. 3C). This level of RegA$_{SMS}$ expression was not significantly different from that of RegA$_{CR}$ (pMUAT7), which displayed 846 ± 31 U under the same conditions (Fig. 3C) ($P > 0.05$; Student's *t* test, two tailed).

To determine if RegA$_{SMS}$ can activate transcription of gene targets, a RegA$_{SMS}$ complementation vector was created by clon-

TABLE 3 Effects of $RegA_{SMS}$ homologues on *C. rodentium* promoters

lacZ construct	β-Galactosidase activity (Miller units) (mean ± SD)[a]		
	No RegA (pACYC184)	RegA homologue	
		$RegA_{CR}$	$RegA_{SMS}$
$regB_{CR}$-*lacZ*	92.11 ± 4.77	4,228.87 ± 428.62	5,074.89 ± 585.63
$regB_{CR}$-*lacZ* mutant	127.35 ± 14.25	476.47 ± 8.15	620.3 ± 17.26
kfc-lacZ	54.85 ± 0.93	5,084.35 ± 662.98	3,035.33 ± 144.49
kfc-lacZ mutant	23.18 ± 2.23	202.5 ± 16.86	297 ± 33.16

[a] β-Galactosidase activity assayed in *E. coli* MC4100 after growth in Luria broth with 45 mM sodium bicarbonate.

ing the $regA_{SMS}$ coding region and 75 bp of its upstream promoter region into pACYC184 to form pAT9. This plasmid, as well as pAT6 and the control plasmid (pACYC184), was then each transformed into *E. coli* MC4100, containing either the *kfc-lacZ* or *regB-lacZ* transcriptional fusion vector (their construction is described in references 54 and 50, respectively) to determine if $RegA_{SMS}$ could activate transcription of *C. rodentium kfc* and *regB*. β-Galactosidase assay data showed that $RegA_{SMS}$ is capable of substantially upregulating $regB_{CR}$ (to 5,074 ± 585 U [mean ± standard deviation {SD}]) and kfc_{CR} (3,035 ± 144 U) to an extent similar to that of $RegA_{CR}$ (4,228 ± 428 U and 5,084 ± 663 U, respectively) (Table 3). These data established that $RegA_{SMS}$ can substitute for $RegA_{CR}$.

We next investigated whether $RegA_{SMS}$ interactions were mediated through the same 5′-GATATAA-3′ DNA consensus binding motif as $RegA_{CR}$ (50). To this end, *E. coli* MC4100 was cotransformed with the pAT9 $RegA_{SMS}$ complementation vector and mutant *regB-lacZ* or *kfc-lacZ* promoter fusions lacking the 5′-GATATAA-3′ motif. β-Galactosidase assay data for these constructs showed that absence of the 5′-GATATAA-3′ motif ablates the ability of $RegA_{SMS}$ to activate the kfc_{CR} and $regB_{CR}$ promoters (Table 3). This indicated that $RegA_{SMS}$ uses the same motif as $RegA_{CR}$ to bind and activate transcription.

$RegA_{SMS}$ can restore partial virulence to *C. rodentium* ICC169 Δ*regA*. Having established that $RegA_{SMS}$ can regulate $RegA_{CR}$ targets, the next question was whether it could complement $RegA_{CR}$ functions in an animal infection model. To investigate this, C57BL/6 mice were infected with *C. rodentium* strain ICC169 or the ICC169 Δ*regA* strain, EMHI, carrying the empty vector (pACYC184) or a complementation vector, pAT9 ($RegA_{SMS}$) or pAT6 ($RegA_{CR}$). Colonization of mice was assessed for up to 15 days. Five mice from each group were sacrificed at the peak of the infectious period to assess colonic colonization and pathology.

Data from these assays confirmed that the colonization ability of the RegA mutant, EMH1, is attenuated compared to that of *C. rodentium* ICC169 (Fig. 4A) (52). When RegA was provided in *trans*, the colonization phenotype was restored, with fecal counts rising to wild-type levels regardless of whether $RegA_{CR}$ or $RegA_{SMS}$ was used (Fig. 4A). In contrast, analysis of colonic colonization revealed ICC169 Δ*regA* complemented with $RegA_{SMS}$ in significantly lower numbers than the same strain complemented with $RegA_{CR}$ (Fig. 4B). These data indicated that fecal counts do not necessarily reflect colonization of the intestinal mucosa *per se*, as fecal counts mainly reflect the numbers of bacteria in the gut lumen.

Mice infected with the $RegA_{SMS}$-complemented strain also differed clinically from those infected with strains containing the native $RegA_{CR}$ homologue. This was manifested by the physical condition of the mice during the experiment, with some infected mice suffering severe effects and requiring culling. The culled mice belonged exclusively to the *C. rodentium* ICC169 group or the ICC169 Δ*regA* strain complemented with native $RegA_{CR}$, with only 50% of the ICC169 and 40% of the pAT6-complemented groups surviving to the end of the experiment (Fig. 5A). In comparison, no clinical signs of disease were evident in mice in the ICC169 Δ*regA* group [EMHI(pACYC184)], with no mice in the group needing to be culled (Fig. 5A). Similarly, mice infected with the ICC169 Δ*regA* strain complemented with $RegA_{SMS}$ [EMHI(pAT9)] showed no evidence of disease, despite having fecal counts equivalent to those of mice infected with the ICC169(pACYC184) or EMHI(pAT6) strain (Fig. 5A).

The clinical findings were supported by histological examination of mouse intestines for colonic hyperplasia on day 10 of infection. This analysis showed that crypt heights did not differ

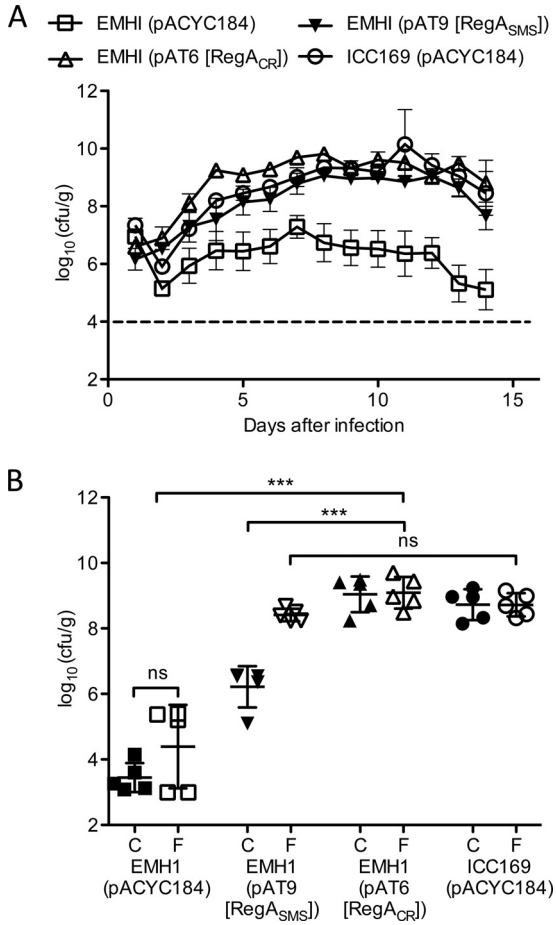

FIG 4 Infection of mice with *C. rodentium* strains. (A) Fecal counts of *C. rodentium* in 4-week-old mice infected with the *C. rodentium* Δ*regA* strain (EMHI) carrying the empty vector pACYC184, pAT9 ($RegA_{SMS}$), or pAT6 ($RegA_{CR}$) or wild-type *C. rodentium* (ICC169) carrying pACYC184. The data were collected for 15 days from 6 to 8 mice in each test or control group and are shown as $\log_{10}$ CFU/g ± standard errors of the mean (SEM), with the limit of detection indicated by the dashed line. (B) Bacterial counts in the colon (C; solid symbols) and feces (F; open symbols) were determined at day 10 of infection for five mice from each group. Colon and fecal counts are shown as $\log_{10}$ CFU/g ± SD. The data for each group were analyzed by ANOVA with the Tukey posttest. ns, $P > 0.05$; ***, $P < 0.001$.

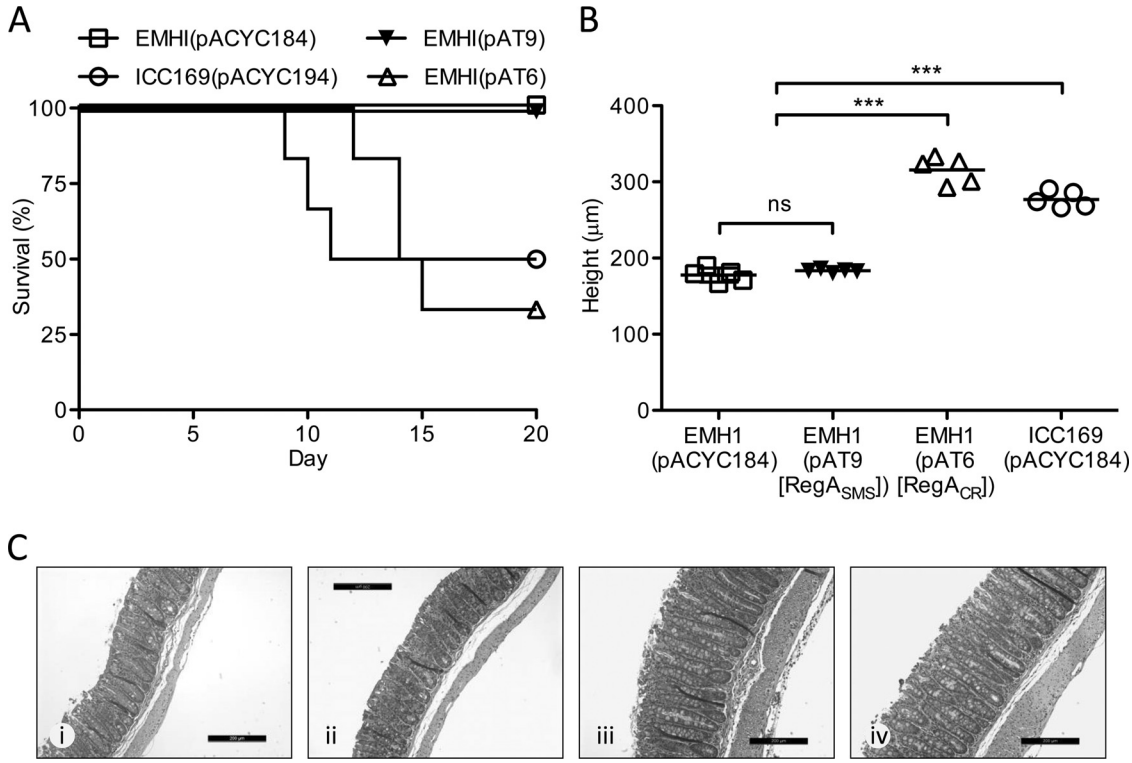

FIG 5 Outcome of infection of mice with *C. rodentium* strains. (A) Kaplan-Meier survival curve of 3-week-old mice infected with the *C. rodentium* Δ*regA* strain (EMHI) carrying the empty vector pACYC184, pAT9 ($RegA_{SMS}$), or pAT6 ($RegA_{CR}$) or wild-type *C. rodentium* (ICC169) carrying pACYC184. The data were collected over 20 days from 6 mice in each group. (B) Colonic-crypt height on day 10 of infection of 4-week old mice, showing the value for each mouse and the mean of each experimental group derived from at least 8 full-height measurements from five mice in each group. Differences in crypt heights were analyzed by ANOVA with the Tukey posttest. ns, $P > 0.05$; ***, $P < 0.001$. (C) Representative pathology of mouse colons on day 10 of infection. The sections show crypt height in mice infected with the *C. rodentium* Δ*regA* strain (EMHI) carrying the empty vector pACYC184 (i), pAT9 ($RegA_{SMS}$) (ii), or pAT6 ($RegA_{CR}$) (iii) or wild-type *C. rodentium* carrying pACYC184 (iv). The sections were stained with hematoxylin and eosin. Bar, 200 μm.

between mice infected with *C. rodentium* $RegA^-$ [EMHI (pACYC184)] or *C. rodentium* $RegA_{SMS}$ (Fig. 5B and C). However, both of these groups showed significantly lower mean crypt heights than mice infected with the $RegA_{CR}$-complemented strain [EMHI(pAT6)] or mice infected with the control strain, ICC169(pACYC184) (Fig. 5B and C). This indicated that, compared to strains with the native $RegA_{CR}$ homologue, the $RegA_{SMS}$-complemented strain did not cause measurable colonic hyperplasia.

$RegA_{CR}$ and $RegA_{SMS}$ differentially regulate the *C. rodentium* genome. To investigate what possible regulatory differences might contribute to the virulence differences *in vivo*, we next examined the ability of $RegA_{SMS}$ to regulate the global $RegA_{CR}$ regulon. To this end, a $RegA_{SMS}$-complemented *C. rodentium* Δ*regA* mutant [EMHI(pAT9)] was grown in the presence of 45 mM bicarbonate, and RNA was harvested at mid-log phase for comparative microarray analysis with similarly grown EMHI strains carrying either the pACYC184 empty vector or the native $RegA_{CR}$ complementation vector, pAT6.

Data from these analyses indicated that $RegA_{SMS}$ and $RegA_{CR}$ regulate housekeeping, stress, and motility genes to similar extents (Table 4). These genes are present in both *C. rodentium* ICC169 and *E. coli* SMS-3-5, suggesting that regulation of these elements is functionally conserved and may be common to all RegA homologues. In contrast, we found some differences in the abilities of $RegA_{SMS}$ and $RegA_{CR}$ to activate virulence genes.

Several virulence genes showed similar regulation by $RegA_{SMS}$ and $RegA_{CR}$, including *sfpA*, *regB*, *adcA*, *kfc*, *aap*, and the *aat* and ROD_41031 to ROD_41051 operons (Table 4). Typically, induction of these genes by $RegA_{CR}$ was higher than that by $RegA_{SMS}$, although the differences did not exceed 1.4-fold (Table 4). However, $RegA_{CR}$ and $RegA_{SMS}$ showed significantly different regulation of LEE PAI and associated effector proteins, with $RegA_{CR}$ inducing between 1.5-fold and 4-fold more production of *grlRA*, *escDFJSTU*, *espABDFGZ*, *sepDLQ*, *cesDD2FT*, *ler*, *tir*, *eae*, *map*, *espIST*, *espM2*, *espN1*, *espN2-1*, *espN2-2*, *nleG1*, *nleG2*, and *nleG8* (Table 4). The biggest differences were seen for the *grlA* and *grlR* genes (4.11- and 4.16-fold more production by $RegA_{CR}$, respectively).

To determine if these differences were caused by the action of $RegA_{SMS}$ at the *grlRA* promoter, the comparative abilities of $RegA_{SMS}$ and $RegA_{CR}$ to activate the *grlRA* promoter (using the *grlRA-1–lacZ*-fusion construct [29] were assessed by β-galactosidase assay. The data from these assays demonstrated that both $RegA_{SMS}$ and $RegA_{CR}$ could induce *grlRA* activation (804.4 ± 75.3 and 2,672.5 ± 431.9 [mean ± S.D.] Miller units of activity, respectively, compared to 289.4 ± 23.8 Miller units for pACYC184). Importantly, however, $RegA_{SMS}$ induced significantly lower levels of activation of the *grlRA* promoter (2.7-fold) than $RegA_{CR}$ (9.2-fold under the same conditions) ($P < 0.0001$; Student's *t* test, two tailed). These data confirmed that $RegA_{SMS}$ was deficient in activation of the *grlRA* promoter, and hence the LEE PAI.

To investigate this further, we compared the secretome profiles of EMH1(pAT6) ($RegA_{CR}$) and EMH1(pAT9) ($RegA_{SMS}$). The results of this analysis showed that LEE expression in $RegA_{SMS}$ is

TABLE 4 *C. rodentium* genes regulated by $RegA_{SMS}$ and $RegA_{CR}$

Gene[a]	Function	Fold induction[b]		Fold difference[c]
		$RegA_{SMS}$	$RegA_{CR}$	
Housekeeping and stress genes				
thrABC	Threonine biosynthesis	2.13	2.19	1.03
tdcAB	Threonine metabolism	−3.21	−3.11	0.97
glpADFKQT	Glycerol metabolism	−3.08	−2.54	0.82
treBC	Trelahose metabolism	−2.78	−3.0	1.08
lamB, *malETK*	Maltose metabolism	−2.81	−2.84	1.01
stpA	Global gene repressor	2.48	1.85	0.75
dnaK	Heat shock protein 70	−2.48	−1.88	1.32
ompFW	Outer membrane protein	−1.92	−2.26	1.31
cspA	Cold shock protein A	−2.49	−2.58	0.97
Motility genes				
flgABCKLMN	Motility	−3.18 to −9.38	−3.15 to −9.29	0.9 to 1.11
flhB	Flagellum biosynthesis protein	−2.07	−2.07	1.0
fliADEFGHJMOSZ	Flagellum cap protein	−2.12 to −6.35	−2.28 to −7.1	0.93 to 1.12
Virulence genes				
aap	Dispersin	78	88.18	1.13
aatABDP	Aat transport system	10.64	12.9	1.21
ROD_41031 to ROD_41051	ABC transporter membrane protein; putative HlyD (*Shewanella woodyi*)	3.11	3.97	1.28
sfpA	Porin protein SfpA (pseudogene)	6.36	7.37	1.16
fimA	Major subunit of the type I fimbriae (pilin)	−3.82	−5.43	1.42
kfcCEFH	Kfc fimbrial operon	5.26	5.25	0.99
adcA	AdcA autotransporter	5.65	6.59	1.17
pagC	PagC homologue	2.85	3.96	1.39
regB	Regulator protein	42.34	45.79	1.08
LEE PAI genes				
espG	R1/R2; T3SS[d] effector protein	NS	2.28	−2.21
LEE4: *espABDF*, *escF*, *sepL*, and ROD29711	LEE4 operon; effector protein	1.45	3.32	2.29
Tir operon: *eae*, *cesT*, *tir*, and *map*	Tir operon; intimin	NS to 2.01	2.92 to 5.01	2.02 to 2.49
LEE3: *cesF*, *sepQ*, and miscellaneous	Chaperone	1.64	3.44	2.09
LEE2: *cesD*, *espJ*, *escZ*, *sepD*, and miscellaneous	LEE2 operon; effector protein	1.46	2.9	1.97
grlAR	Global regulators of LEE	2.34	9.67	4.13
LEE1: *escSTU*, *orf345*, and *ler*	LEE1 operon; structural protein	1.86	4.61	2.41
espN1, *espM2*, *espS*, *espT*, *espX7*, *espI* (*nleA*), and *nleG1*	Effector proteins	NS to 1.59	2.04 to 4.76	2.04 to 4.03

[a] ORF and gene names and functions are as annotated in the *C. rodentium* genome in GenBank (accession number FN543502). In some cases, the annotation and function have been determined with reference to BLAST searches and EcoCyc (http://www.ecocyc.org). ORFs are shown in numerical order and grouped by functional category.
[b] $RegA_{SMS}$ or $RegA_{CR}$ activation or repression (−) of the target, relative to the pACYC184 empty-vector control condition, for assays conducted in LB broth with 45 mM bicarbonate. Data were included only where the average magnitude of the change exceeded 2-fold for either condition and the adjusted *P* value for all assays was <0.05. NS, transcription was insufficient to generate a signal on the array. Note that for gene groups, the change presented is an average of all induction levels; individual levels varied.
[c] The fold difference between $RegA_{SMS}$ and $RegA_{CR}$ induction or repression levels, calculated as the fold change in transcription induced/repressed by $RegA_{CR}$ divided by the fold change induced/repressed by $RegA_{SMS}$. Where $RegA_{SMS}$ induction was not calculated, the value was derived by direct comparison of the $RegA_{SMS}$ and $RegA_{CR}$ data generated by the Agilent Feature Extraction software.
[d] T3SS, type 3 secretion system.

reduced compared to that of $RegA_{CR}$ (see Fig. S4 in the supplemental material). Together, our findings suggest that *C. rodentium* Δ*regA* complemented with $RegA_{SMS}$ showed altered colonization and reduced virulence for mice compared to the same strain complemented with $RegA_{CR}$ due to the inability of $RegA_{SMS}$ to fully activate the LEE PAI in *C. rodentium*.

C. rodentium requires the LEE PAI to adhere to the intestinal mucosa and for the production of colonic hyperplasia and the severe illness leading to death in some animals. The inability of $RegA_{SMS}$ to fully activate LEE expression via GrlA would lead to fewer tightly adherent *C. rodentium* cells on the epithelium and the absence of A/E lesions, resulting in infections without serious consequences. Our findings corroborate previous studies that indicated that loss of the LEE PAI alters the colonization dynamics of *C. rodentium*, relocating the bacteria to the lumen of the gut (79). Our study also demonstrates the importance of $RegA_{CR}$ for the expression of the LEE and the full virulence of *C. rodentium* and suggests that the horizontal acquisition of *regA* is likely to have aided the development of *C. rodentium* as a pathogen.

Previous analyses of the *C. rodentium* genome have indicated that the pathogen is adapting to a new niche in the mouse, with high levels of horizontal gene influx and inactivation of core genome components in favor of transferred DNA (78). Many of the horizontally acquired elements are virulence factors, including the

LEE PAI (58), which are typically prone to repression by H-NS/StpA (80–84). The acquisition of *regA* would have provided *C. rodentium* with an AraC/XylS regulator capable of displacing or disrupting H-NS/StpA repression in a nonspecific manner (85). If acquisition of the *regA* locus preceded the acquisition of the LEE PAI, promiscuous derepression of H-NS and acquisition of the *adcA*- and *kfc*-encoded adhesins may have been sufficient to induce asymptomatic colonization of the mouse intestinal lumen, placing *C. rodentium* in a "prepathogenic" state (8). Promiscuous H-NS derepression by $RegA_{CR}$ may also have played a role in the activation of the LEE PAI once it had been acquired, leading to the development of the $RegA_{CR}$-LEE PAI interactions that are necessary for the full virulence of *C. rodentium*. These interactions are strongly influenced by species- and strain-specific sequences in $RegA_{CR}$.

Our data also imply that the evolution of LEE-external regulators is an *ad hoc* process in which cells coopt regulators to facilitate the expression of LEE. In the case of *C. rodentium*, the regulator coopted was $RegA_{CR}$, which was likely originally a regulator of Kfc fimbriae and the AdcA autotransporter, encoded in the *regA* locus. Interestingly, this is also seen in connection with other regulators involved in LEE PAI regulation. For example, in typical EPEC strains, the LEE PAI is externally activated by PerC and indirectly by PerA, which also regulates the *bfp* structural genes, encoding bundle-forming pili (86, 87). In EHEC, activation of the LEE PAI is linked to the Pch regulators. While they do not specifically affect adhesins other than those encoded by LEE, the *pch* genes are regulated by the LysR homologue LrhA (88), which is also involved in the repression of motility, chemotaxis, and the expression of type I fimbriae (89) and enterohemolysin (90). This suggests that, while the coopting of external LEE PAI regulators may be *ad hoc* and opportunistic, it may also selectively favor regulators that already mediate or enhance colonization.

Conclusions. The findings of this study have implications for the evolution of both virulence and transcriptional regulators. Most importantly, this work indicates that pathogen evolution is not solely a consequence of the horizontal transfer of virulence elements. While the transfer of virulence determinants can cause the rapid and dramatic emergence of pathogenic strains, our study showed that these elements may also need to be embedded in the existing regulatory circuits of the recipient strain to permit their expression when needed. Our analysis also indicated that this process may require strain-specific sequence changes on the part of the regulatory circuit, as well, and that the capacity for mutability of the regulator function in response to host strain and selective pressures may affect the development of pathogenicity itself. Finally, our findings suggest that these processes may also occur with horizontally transferred regulatory elements, with RegA likely to have been transferred into *C. rodentium* and its regulatory functions then coopted for the colonization of mice. Collectively these studies suggest that pathogenicity may evolve, not only from the horizontal acquisition of virulence genes, but through the horizontal transfer and evolutionary coopting of nonpathogenic elements, representing a mechanism by which bacterial species can prime themselves for the exploitation of new niches and hasten and facilitate their development into pathogens if this improves their chances of survival.

ACKNOWLEDGMENTS

We are grateful to David Gordon of the Australian National University and Jacques Ravel of the University of Maryland for their generous provision of bacterial strains and unpublished data. We also thank Hans Steinsland, University of Bergen, for useful discussions about MLST.

This work was supported by grants from the Australian National Health and Medical Research Council (NHMRC).

REFERENCES

1. **Burrus V, Waldor MK.** 2004. Shaping bacterial genomes with integrative and conjugative elements. Res Microbiol **155:**376–386. http://dx.doi.org/10.1016/j.resmic.2004.01.012.
2. **Juhas M, Van Der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW.** 2009. Genomic islands: tools of bacterial horizontal gene transfer and evolution. FEMS Microbiol Rev **33:**376–393. http://dx.doi.org/10.1111/j.1574-6976.2008.00136.x.
3. **Dobrindt U, Hochhut B, Hentschel U, Hacker J.** 2004. Genomic islands in pathogenic and environmental microorganisms. Nat Rev Microbiol **2:**414–424. http://dx.doi.org/10.1038/nrmicro884.
4. **Ochman H, Lawrence JG, Groisman EA.** 2000. Lateral gene transfer and the nature of bacterial innovation. Nature **405:**299–304. http://dx.doi.org/10.1038/35012500.
5. **Ziebuhr W, Ohlsen K, Karch H, Korhonen T, Hacker J.** 1999. Evolution of bacterial pathogenesis. Cell Mol Life Sci **56:**719–728. http://dx.doi.org/10.1007/s000180050018.
6. **Groisman EA, Ochman H.** 1996. Pathogenicity islands: bacterial evolution in quantum leaps. Cell **87:**791–794. http://dx.doi.org/10.1016/S0092-8674(00)81985-6.
7. **Hacker J, Kaper JB.** 2000. Pathogenicity islands and the evolution of microbes. Annu Rev Microbiol **54:**641–679. http://dx.doi.org/10.1146/annurev.micro.54.1.641.
8. **Jackson RW, Vinatzer B, Arnold DL, Dorus S, Murillo J.** 2011. The influence of the accessory genome on bacterial pathogen evolution. Mob Genet Elements **1:**55–65. http://dx.doi.org/10.4161/mge.1.1.16432.
9. **Jackson RW, Johnson LJ, Clarke SR, Arnold DL.** 2011. Bacterial pathogen evolution: breaking news. Trends Genet **27:**32–40. http://dx.doi.org/10.1016/j.tig.2010.10.001.
10. **Wren BW.** 2003. The Yersiniae—a model genus to study the rapid evolution of bacterial pathogens. Nat Rev Microbiol **1:**55–64. http://dx.doi.org/10.1038/nrmicro730.
11. **Mellmann A, Harmsen D, Cummings CA, Zentz EB, Leopold SR, Rico A, Prior K, Szczepanowski R, Ji Y, Zhang W, McLaughlin SF, Henkhaus JK, Leopold B, Bielaszewska M, Prager R, Brzoska PM, Moore RL, Guenther S, Rothberg JM, Karch H.** 2011. Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. PLoS One **6:**e22751. http://dx.doi.org/10.1371/journal.pone.0022751.
12. **Frank C, Werber D, Cramer JP, Askar M, Faber M, an der Heiden M, Bernard H, Fruth A, Prager R, Spode A, Wadl M, Zoufaly A, Jordan S, Kemper MJ, Follin P, Müller L, King LA, Rosner B, Buchholz U, Stark K, Krause G.** 2011. Epidemic profile of Shiga-toxin-producing Escherichia coli O104:H4 outbreak in Germany. N Engl J Med **365:**1771–1780. http://dx.doi.org/10.1056/NEJMoa1106483.
13. **Rasko DA, Webster DR, Sahl JW, Bashir A, Boisen N, Scheutz F, Paxinos EE, Sebra R, Chin C-S, Iliopoulos D, Klammer A, Peluso P, Lee L, Kislyuk AO, Bullard J, Kasarskis A, Wang S, Eid J, Rank D, Redman JC, Steyert SR, Frimodt-Møller J, Struve C, Petersen AM, Krogfelt KA, Nataro JP, Schadt EE, Waldor MK.** 2011. Origins of the *E. coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. N Engl J Med **365:**709–717. http://dx.doi.org/10.1056/NEJMoa1106920.
14. **Rohde H, Qin J, Cui Y, Li D, Loman NJ, Hentschke M, Chen W, Pu F, Peng Y, Li J, Xi F, Li S, Li Y, Zhang Z, Yang X, Zhao M, Wang P, Guan Y, Cen Z, Zhao X, Christner M, Kobbe R, Loos S, Oh J, Yang L, Danchin A, Gao GF, Song Y, Li Y, Yang H, Wang J, Xu J, Pallen MJ, Wang J, Aepfelbacher M, Yang R.** 2011. Open-source genomic analysis of Shiga-toxin-producing E. coli O104:H4. N Engl J Med **365:**718–724. http://dx.doi.org/10.1056/NEJMoa1107643.
15. **Deng W, Li Y, Vallance BA, Finlay BB.** 2001. Locus of enterocyte effacement from *Citrobacter rodentium*: sequence analysis and evidence for horizontal transfer among attaching and effacing pathogens. Infect Immun **69:**6323–6335. http://dx.doi.org/10.1128/IAI.69.10.6323-6335.2001.
16. **Jores J, Rumer L, Wieler LH.** 2004. Impact of the locus of enterocyte effacement pathogenicity island on the evolution of pathogenic *Escherichia coli*. Int J Med Microbiol **294:**103–113. http://dx.doi.org/10.1016/j.ijmm.2004.06.024.

17. **Konczy P, Ziebell K, Mascarenhas M, Choi A, Michaud C, Kropinski AM, Whittam TS, Wickham M, Finlay B, Karmali MA.** 2008. Genomic O island 122, locus for enterocyte effacement, and the evolution of virulent verocytotoxin-producing *Escherichia coli*. J Bacteriol **190:**5832–5840. http://dx.doi.org/10.1128/JB.00480-08.
18. **Schmidt MA.** 2010. LEEways: tales of EPEC, ATEC and EHEC. Cell Microbiol **12:**1544–1552. http://dx.doi.org/10.1111/j.1462-5822.2010.01518.x.
19. **Hyma KE, Lacher DW, Nelson AM, Bumbaugh AC, Janda JM, Strockbine NA, Young VB, Whittam TS.** 2005. Evolutionary genetics of a new pathogenic *Escherichia* species: *Escherichia albertii* and related *Shigella boydii* strains. J Bacteriol **187:**619–628. http://dx.doi.org/10.1128/JB.187.2.619-628.2005.
20. **Chandry PS, Gladman S, Moore SC, Seemann T, Crandall KA, Fegan N.** 2012. A genomic island in *Salmonella enterica ssp. salamae* provides new insights on the genealogy of the locus of enterocyte effacement. PLoS One **7:**e41615. http://dx.doi.org/10.1371/journal.pone.0041615.
21. **O'Gorman LE, Krejany EO, Bennett-Wood VR, Robins-Browne RM.** 1996. Transfer of attaching and effacing capacity from a strain of enteropathogenic *Escherichia coli* to E. coli K-12. Microbiol Res **151:**379–385. http://dx.doi.org/10.1016/S0944-5013(96)80007-3.
22. **Elliott SJ, Yu J, Kaper JB.** 1999. The cloned locus of enterocyte effacement from enterohemorrhagic *Escherichia coli* O157:H7 is unable to confer the attaching and effacing phenotype upon *E. coli* K-12. Infect Immun **67:**4260–4263.
23. **Beutin L.** 2006. Emerging enterohaemorrhagic *Escherichia coli*: causes and effects of the rise of a human pathogen. J Vet Med B Infect Dis Vet Public Health **53:**299–305. http://dx.doi.org/10.1111/j.1439-0450.2006.00968.x.
24. **Ferens WA, Hovde CJ.** 2011. *Escherichia coli* O157:H7: animal reservoir and sources of human infection. Foodborne Pathog Dis **8:**465–487. http://dx.doi.org/10.1089/fpd.2010.0673.
25. **Brandal LT, Sekse C, Lindstedt B-A, Sunde M, Løbersli I, Urdahl AM, Kapperud G.** 2012. Norwegian sheep are an important reservoir for human-pathogenic *Escherichia coli* O26:H11. Appl Environ Microbiol **78:**4083–4091. http://dx.doi.org/10.1128/AEM.00186-12.
26. **Oh J-Y, Kang M-S, An B-K, Shin E-G, Kim M-J, Kim Y-J, Kwon Y-K.** 2012. Prevalence and characteristics of intimin-producing *Escherichia coli* strains isolated from healthy chickens in Korea. Poult Sci **91:**2438–2443. http://dx.doi.org/10.3382/ps.2012-02301.
27. **Barba J, Bustamante VH, Flores-Valdez MA, Deng W, Finlay BB, Puente JL.** 2005. A positive regulatory loop controls expression of the locus of enterocyte effacement-encoded regulators Ler and GrlA. J Bacteriol **187:**7918–7930. http://dx.doi.org/10.1128/JB.187.23.7918-7930.2005.
28. **Bustamante VH, Villalba MI, García-Angulo VA, Vázquez A, Martínez LC, Jiménez R, Puente JL.** 2011. PerC and GrlA independently regulate Ler expression in enteropathogenic *Escherichia coli*. Mol Microbiol **82:**398–415. http://dx.doi.org/10.1111/j.1365-2958.2011.07819.x.
29. **Tauschek M, Yang J, Hocking D, Azzopardi K, Tan A, Hart E, Praszkier J, Robins-Browne RM.** 2010. Transcriptional analysis of the *grlRA* virulence operon from *Citrobacter rodentium*. J Bacteriol **192:**3722–3734. http://dx.doi.org/10.1128/JB.01540-09.
30. **Ibarra JA, Villalba MI, Puente JL.** 2003. Identification of the DNA binding sites of PerA, the transcriptional activator of the *bfp* and *per* operons in enteropathogenic *Escherichia coli*. J Bacteriol **185:**2835–2847. http://dx.doi.org/10.1128/JB.185.9.2835-2847.2003.
31. **Martinez-Laguna Y, Calva E, Puente JL.** 1999. Autoactivation and environmental regulation of *bfpT* expression, the gene coding for the transcriptional activator of *bfpA* in enteropathogenic *Escherichia coli*. Mol Microbiol **33:**153–166. http://dx.doi.org/10.1046/j.1365-2958.1999.01460.x.
32. **Porter ME, Mitchell P, Roe AJ, Free A, Smith DGE, Gally DL.** 2004. Direct and indirect transcriptional activation of virulence genes by an AraC-like protein, PerA from enteropathogenic *Escherichia coli*. Mol Microbiol **54:**1117–1133. http://dx.doi.org/10.1111/j.1365-2958.2004.04333.x.
33. **Bodero MD, Pilonieta MC, Munson GP.** 2007. Repression of the inner membrane lipoprotein NlpA by Rns in enterotoxigenic *Escherichia coli*. J Bacteriol **189:**1627–1632. http://dx.doi.org/10.1128/JB.01714-06.
34. **Caron J, Coffield LM, Scott JR.** 1989. A plasmid-encoded regulatory gene, *rns*, required for expression of the CS1 and CS2 adhesins of enterotoxigenic *Escherichia coli*. Proc Natl Acad Sci U S A **86:**963–967. http://dx.doi.org/10.1073/pnas.86.3.963.
35. **Froehlich B, Husmann L, Caron J, Scott JR.** 1994. Regulation of *rns*, a positive regulatory factor for pili of enterotoxigenic *Escherichia coli*. J Bacteriol **176:**5385–5392.
36. **Munson GP, Scott JR.** 1999. Binding site recognition by Rns, a virulence regulator in the AraC family. J Bacteriol **181:**2110–2117.
37. **Dudley EG, Thomson NR, Parkhill J, Morin NP, Nataro JP.** 2006. Proteomic and microarray characterization of the AggR regulon identifies a *pheU* pathogenicity island in enteroaggregative *Escherichia coli*. Mol Microbiol **61:**1267–1282. http://dx.doi.org/10.1111/j.1365-2958.2006.05281.x.
38. **Morin N, Santiago AE, Ernst RK, Guillot SJ, Nataro JP.** 2013. Characterization of the AggR regulon in enteroaggregative *E. coli*. Infect Immun **81:**122–132. http://dx.doi.org/10.1128/IAI.00676-12.
39. **Morin N, Tirling C, Ivison SM, Kaur AP, Nataro JP, Steiner TS.** 2010. Autoactivation of the AggR regulator of enteroaggregative *Escherichia coli in vitro* and *in vivo*. FEMS Immunol Med Microbiol **58:**344–355. http://dx.doi.org/10.1111/j.1574-695X.2010.00645.x.
40. **Nataro JP, Yikang D, Yingkang D, Walker K.** 1994. AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli*. J Bacteriol **176:**4691–4699.
41. **Krukonis ES, Yu RR, Dirita VJ.** 2000. The *Vibrio cholerae* ToxR/TcpP/ToxT virulence cascade: distinct roles for two membrane-localized transcriptional activators on a single promoter. Mol Microbiol **38:**67–84. http://dx.doi.org/10.1046/j.1365-2958.2000.02111.x.
42. **Schuhmacher DA, Klose KE.** 1999. Environmental signals modulate ToxT-dependent virulence factor expression in *Vibrio cholerae*. J Bacteriol **181:**1508–1514.
43. **Withey JH, DiRita VJ.** 2005. *Vibrio cholerae* ToxT independently activates the divergently transcribed *aldA* and *tagA* genes. J Bacteriol **187:**7890–7900. http://dx.doi.org/10.1128/JB.187.23.7890-7900.2005.
44. **Yu RR, DiRita VJ.** 2002. Regulation of gene expression in *Vibrio cholerae* by ToxT involves both antirepression and RNA polymerase stimulation. Mol Microbiol **43:**119–134. http://dx.doi.org/10.1046/j.1365-2958.2002.02721.x.
45. **Falconi M, Colonna B, Prosseda G, Micheli G, Gualerzi CO.** 1998. Thermoregulation of *Shigella* and *Escherichia coli* EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of *virF* promoter to transcriptional repressor H-NS. EMBO J **17:**7033–7043.
46. **Sakai T.** 1988. Expression of four virulence antigens of *Shigella flexneri* is positively regulated at the transcriptional level by the 30 kilodalton *virF* protein. Mol Microbiol **2:**589. http://dx.doi.org/10.1111/j.1365-2958.1988.tb00067.x.
47. **Tran CN, Giangrossi M, Prosseda G, Brandi A, Di Martino ML, Colonna B, Falconi M.** 2011. A multifactor regulatory circuit involving H-NS, VirF and an antisense RNA modulates transcription of the virulence gene *icsA* of *Shigella flexneri*. Nucleic Acids Res **39:**8122–8134. http://dx.doi.org/10.1093/nar/gkr521.
48. **Withey JH, DiRita VJ.** 2006. The toxbox: specific DNA sequence requirements for activation of *Vibrio cholerae* virulence genes by ToxT. Mol Microbiol **59:**1779–1789. http://dx.doi.org/10.1111/j.1365-2958.2006.05053.x.
49. **Thomas VJ, Collins CM.** 1999. Identification of UreR binding sites in the Enterobacteriaceae plasmid-encoded and *Proteus mirabilis* urease gene operons. Mol Microbiol **31:**1417–1428. http://dx.doi.org/10.1046/j.1365-2958.1999.01283.x.
50. **Tan A, Yang J, Tauschek M, Praszkier J, Robins-Browne RM.** 2011. Autogenous transcriptional regulation of the *regA* gene, encoding an AraC-like, essential virulence regulator in *Citrobacter rodentium*. J Bacteriol **193:**1777–1782. http://dx.doi.org/10.1128/JB.01224-10.
51. **Poore CA, Mobley HLT.** 2003. Differential regulation of the *Proteus mirabilis* urease gene cluster by UreR and H-NS. Microbiology **149:**3383–3394. http://dx.doi.org/10.1099/mic.0.26624-0.
52. **Hart E, Yang J, Tauschek M, Kelly M, Wakefield MJ, Frankel G, Hartland EL, Robins-Browne RM.** 2008. RegA, an AraC-like protein, is a global transcriptional regulator that controls virulence gene expression in *Citrobacter rodentium*. Infect Immun **76:**5247–5256. http://dx.doi.org/10.1128/IAI.00770-08.
53. **Kelly M, Hart E, Mundy R, Marches O, Wiles S, Badea L, Luck S, Tauschek M, Frankel G, Robins-Browne RM, Hartland EL.** 2006. Essential role of the type III secretion system effector NleB in colonization of mice by *Citrobacter rodentium*. Infect Immun **74:**2328–2337. http://dx.doi.org/10.1128/IAI.74.4.2328-2337.2006.
54. **Yang J, Hart E, Tauschek M, Price GD, Hartland EL, Strugnell RA,**

Robins-Browne RM. 2008. Bicarbonate-mediated transcriptional activation of divergent operons by the virulence regulatory protein, RegA, from *Citrobacter rodentium*. Mol Microbiol **68:**314–327. http://dx.doi.org/10.1111/j.1365-2958.2008.06171.x.
55. **Yang J, Dogovski C, Hocking D, Tauschek M, Perugini M, Robins-Browne RM.** 2009. Bicarbonate-mediated stimulation of RegA, the global virulence regulator from *Citrobacter rodentium*. J Mol Biol **394:**591–599. http://dx.doi.org/10.1016/j.jmb.2009.10.033.
56. **Yang J, Tauschek M, Hart E, Hartland EL, Robins-Browne RM.** 2010. Virulence regulation in *Citrobacter rodentium*: the art of timing. Microb Biotechnol **3:**259–268. http://dx.doi.org/10.1111/j.1751-7915.2009.00114.x.
57. **Srikhanta Y, Hocking DM, Praszkier J, Wakefield MJ, Robins-Browne RM, Yang J, Tauschek M.** 2013. The RegR virulence regulon of rabbit specific enteropathogenic *Escherichia coli* strain E22. Infect Immun **81:**1078–1089. http://dx.doi.org/10.1128/IAI.01325-12.
58. **Petty NK, Bulgin R, Crepin VF, Cerdeno-Tarraga AM, Schroeder GN, Quail MA, Lennard N, Corton C, Barron A, Clark L, Toribio AL, Parkhill J, Dougan G, Frankel G, Thomson NR.** 2010. The *Citrobacter rodentium* genome sequence reveals convergent evolution with human pathogenic *Escherichia coli*. J Bacteriol **192:**525–538. http://dx.doi.org/10.1128/JB.01144-09.
59. **Casadaban MJ.** 1976. Transposition and fusion of the *lac* genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. J Mol Biol **104:**541–555. http://dx.doi.org/10.1016/0022-2836(76)90119-4.
60. **Fricke WF, Wright MS, Lindell AH, Harkins DM, Baker-Austin C, Ravel J, Stepanauskas R.** 2008. Insights into the environmental resistance gene pool from the genome sequence of the multidrug-resistant environmental isolate *Escherichia coli* SMS-3-5. J Bacteriol **190:**6779–6794. http://dx.doi.org/10.1128/JB.00661-08.
61. **Praszkier J, Wilson IW, Pittard AJ.** 1992. Mutations affecting translational coupling between the *rep* genes of an IncB miniplasmid. J Bacteriol **174:**2376–2383.
62. **Lawley B, Pittard AJ.** 1994. Regulation of *aroL* expression by TyrR protein and Trp repressor in *Escherichia coli* K-12. J Bacteriol **176:**6921–6930.
63. **Campanella J, Bitincka L, Smalley J.** 2003. MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences. BMC Bioinformatics **4:**29. http://dx.doi.org/10.1186/1471-2105-4-29.
64. **Steinsland H, Lacher DW, Sommerfelt H, Whittam TS.** 2010. Ancestral lineages of human enterotoxigenic *Escherichia coli*. J Clin Microbiol **48:**2916–2924. http://dx.doi.org/10.1128/JCM.02432-09.
65. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol **28:**2731–2739. http://dx.doi.org/10.1093/molbev/msr121.
66. **Sullivan MJ, Petty NK, Beatson SA.** 2011. Easyfig: a genome comparison visualizer. Bioinformatics **27:**1009–1010. http://dx.doi.org/10.1093/bioinformatics/btr039.
67. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945. http://dx.doi.org/10.1093/bioinformatics/16.10.944.
68. **Miller JH.** 1974. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
69. **Smyth GK.** 12 February 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol **3:**Article3. http://dx.doi.org/10.2202/1544-6115.1027.
70. **Smyth GK.** 2005. Limma: linear models for microarray data, p 397–420. *In* Gentleman VC, Dudoit S, Irizarry R, Huber W (ed), Bioinformatics and computational biology solutions using R and Bioconductor. Springer, New York, NY.
71. **Smyth GK, Speed T.** 2003. Normalization of cDNA microarray data. Methods **31:**265–273. http://dx.doi.org/10.1016/S1046-2023(03)00155-5.
72. **Edgar R, Domrachev M, Lash AE.** 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res **30:**207–210. http://dx.doi.org/10.1093/nar/30.1.207.
73. **Tennant S, Tauschek M, Azzopardi K, Bigham A, Bennett-Wood V, Hartland E, Qi W, Whittam T, Robins-Browne R.** 2009. Characterisation of atypical enteropathogenic *E. coli* strains of clinical origin. BMC Microbiol **9:**117. http://dx.doi.org/10.1186/1471-2180-9-117.
74. **Retchless AC, Lawrence JG.** 2010. Phylogenetic incongruence arising from fragmented speciation in enteric bacteria. Proc Natl Acad Sci U S A **107:**11453–11458. http://dx.doi.org/10.1073/pnas.1001291107.
75. **Tomich M, Planet PJ, Figurski DH.** 2007. The tad locus: postcards from the widespread colonization island. Nat Rev Microbiol **5:**363–375. http://dx.doi.org/10.1038/nrmicro1636.
76. **Feulner G, Gray JA, Kirschman JA, Lehner AF, Sadosky AB, Vlazny DA, Zhang J, Zhao S, Hill CW.** 1990. Structure of the *rhsA* locus from Escherichia coli K-12 and comparison of *rhsA* with other members of the *rhs* multigene family. J Bacteriol **172:**446–456.
77. **Lin R-J, Capage M, Hill CW.** 1984. A repetitive DNA sequence, rhs, responsible for duplications within the *Escherichia coli* K-12 chromosome. J Mol Biol **177:**1–18. http://dx.doi.org/10.1016/0022-2836(84)90054-8.
78. **Petty NK, Feltwell T, Pickard D, Clare S, Toribio AL, Fookes M, Roberts K, Monson R, Nair S, Kingsley RA, Bulgin R, Wiles S, Goulding D, Keane T, Corton C, Lennard N, Harris D, Willey D, Rance R, Yu L, Choudhary JS, Churcher C, Quail MA, Parkhill J, Frankel G, Dougan G, Salmond GPC, Thomson NR.** 2011. *Citrobacter rodentium* is an unstable pathogen showing evidence of significant genomic flux. PLoS Pathog **7:**e1002018. http://dx.doi.org/10.1371/journal.ppat.1002018.
79. **Kamada N, Kim Y-G, Sham HP, Vallance BA, Puente JL, Martens EC, Núñez G.** 2012. Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. Science **336:**1325–1329. http://dx.doi.org/10.1126/science.1222195.
80. **Dorman CJ.** 2004. H-NS: a universal regulator for a dynamic genome. Nat Rev Microbiol **2:**391–400. http://dx.doi.org/10.1038/nrmicro883.
81. **Dorman CJ.** 2007. H-NS, the genome sentinel. Nat Rev Microbiol **5:**157–161. http://dx.doi.org/10.1038/nrmicro1598.
82. **Navarre WW, McClelland M, Libby SJ, Fang FC.** 2007. Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. Genes Dev **21:**1456–1471. http://dx.doi.org/10.1101/gad.1543107.
83. **Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC.** 2006. Selective silencing of foreign DNA with Low GC content by the H-NS protein in *Salmonella*. Science **313:**236–238. http://dx.doi.org/10.1126/science.1128794.
84. **Oshima T, Ishikawa S, Kurokawa K, Aiba H, Ogasawara N.** 2006. *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. DNA Res **13:**141–153. http://dx.doi.org/10.1093/dnares/dsl009.
85. **Stoebel DM, Free A, Dorman CJ.** 2008. Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. Microbiology **154:**2533–2545. http://dx.doi.org/10.1099/mic.0.2008/020693-0.
86. **Donnenberg MS, Giron JA, Nataro JP, Kaper JB.** 1992. A plasmid-encoded type IV fimbrial gene of enteropathogenic *Escherichia coli* associated with localized adherence. Mol Microbiol **6:**3427–3437. http://dx.doi.org/10.1111/j.1365-2958.1992.tb02210.x.
87. **Ramboarina S, Fernandes PJ, Daniell S, Islam S, Simpson P, Frankel G, Booy F, Donnenberg MS, Matthews S.** 2005. Structure of the bundle-forming pilus from enteropathogenic *Escherichia coli*. J Biol Chem **280:**40252–40260. http://dx.doi.org/10.1074/jbc.M508099200.
88. **Honda N, Iyoda S, Yamamoto S, Terajima J, Watanabe H.** 2009. LrhA positively controls the expression of the locus of enterocyte effacement genes in enterohemorrhagic *Escherichia coli* by differential regulation of their master regulators PchA and PchB. Mol Microbiol **74:**1393–1341. http://dx.doi.org/10.1111/j.1365-2958.2009.06937.x.
89. **Lehnen D, Blumer C, Polen T, Wackwitz B, Wendisch VF, Unden G.** 2002. LrhA as a new transcriptional key regulator of flagella, motility and chemotaxis genes in *Escherichia coli*. Mol Microbiol **45:**521–532. http://dx.doi.org/10.1046/j.1365-2958.2002.03032.x.
90. **Iyoda S, Honda N, Saitoh T, Shimuta K, Terajima J, Watanabe H, Ohnishi M.** 2011. Coordinate control of the locus of enterocyte effacement and enterohemolysin genes by multiple common virulence regulators in enterohemorrhagic *Escherichia coli*. Infect Immun **79:**4628–4637. http://dx.doi.org/10.1128/IAI.05023-11.
91. **Livingstone CD, Barton GJ.** 1993. Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. Comput Appl Biosci **9:**745–756.