

# Global Regulation by Horizontally Transferred Regulators Establishes the Pathogenicity of *Escherichia coli*

Hiroyuki ABE<sup>1</sup>, Akira MIYAHARA<sup>1</sup>, Taku OSHIMA<sup>2</sup>, Kosuke TASHIRO<sup>3</sup>, Yoshitoshi OGURA<sup>4</sup>, Satoru KUHARA<sup>3</sup>, Naotake OGASAWARA<sup>2</sup>, Tetsuya HAYASHI<sup>4</sup>, and Toru TOBE<sup>1,\*</sup>

Division of Applied Bacteriology, Graduate School of Medicine, Osaka University, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan<sup>1</sup>; Graduate School of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0101, Japan<sup>2</sup>; Graduate School of Genetic Resources Technology, Kyusyu University, 6-10-1 Hakozaki, Fukuoka 812-8581, Japan<sup>3</sup> and Department of Microbiology, School of Medicine, Miyazaki University, 5200 Kiyotake, Miyazaki 899-1692, Japan<sup>4</sup>

(Received 4 September 2007; accepted 11 December 2007; published online 25 January 2008)

## Abstract

**Enterohemorrhagic *Escherichia coli* is an emerging pathogen that causes diarrhea and hemolytic uremic syndrome. Much of the genomic information that affects virulence is acquired by horizontal transfer. Genes necessary for attaching and effacing lesions are located in the locus for enterocyte effacement (LEE) pathogenicity island. LEE gene transcription is positively regulated by Ler, which is also encoded by the LEE, and by Pch regulators, which are encoded at other loci. Here we identified genes whose transcription profiles were similar to those of the LEE genes, by comparing the effects of altering *ler* and *pch* transcript levels. We assigned these genes into two classes, according to their transcription profiles. By determining the binding profiles for Ler and Pch, we showed that both were involved in regulating one class of genes, but only Pch was involved in regulating the other class. Binding sites were found in the coding region as well as the promoter region of regulated genes, which include genes common to K12 strains as well as O157-specific genes, suggesting that both act as a global regulator. These results indicate that Ler and Pch orchestrate the transcription of virulence genes, which are captured by horizontal transfer and scattered throughout the chromosome.**

**Key words:** transcriptional regulation; virulence gene; A/E pathogen; DNA-binding protein; captured genome

## 1. Introduction

Pathogenic bacteria are distinct from related non-pathogenic strains in that they possess and express virulence genes. The major avenue for the acquisition of virulence genes is the horizontal transfer of DNA segments.<sup>1</sup> In many pathogenic bacteria, multiple virulence factors function to achieve successful infection and population bursts in niches by protecting bacteria from host defense systems. Multiple virulence

genes in different chromosomal loci must often be simultaneously expressed during the various stages of infection. Thus, the integration of the acquired genes, which are scattered throughout the chromosome, into the gene-regulatory systems is important for establishing pathogenic strains. Because the repertoire of virulence traits is often variable among the strains of the same species, bacterial pathogenicity is a dynamic process involving the ongoing addition of novel virulence genes. Therefore, plasticity that allows newly transferred virulence genes to be adopted must also be a characteristic of the virulence-regulatory system. To understand the emergence and evolution of pathogenic bacteria, therefore, the molecular basis of virulence gene regulation must be elucidated.

---

Edited by Katsumi Isono.

\* To whom correspondence should be addressed. Tel. +81 6-6879-3301. Fax. +81 6-6879-3309. E-mail: torutobe@bact.med.osaka-u.ac.jp

© The Author 2008. Kazusa DNA Research Institute.

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact journals.permissions@oxfordjournals.org

Enterohemorrhagic *Escherichia coli* (EHEC) causes hemorrhagic colitis and hemolytic uremic syndrome.<sup>2</sup> A comparison of the genomic sequences of EHEC O157:H7 Sakai strain with that of the non-pathogenic *E. coli* strain K-12 revealed many EHEC O157-specific loci (called S-loops), most of which contain gene clusters.<sup>3</sup> One of the virulence-associated loci is the locus for enterocyte effacement (LEE), which encodes a type III secretion system (T3SS), type III secreted effectors, and adhesin. These virulence factors are necessary for the induction of attaching and effacing (A/E) lesions in epithelial cells, which are characterized by a rearrangement of the cytoskeleton and the destruction of microvilli.<sup>4,5</sup> In addition to the LEE, gene-encoding effectors for T3SS can be found elsewhere on the chromosome,<sup>6</sup> mostly at the end of lambda-doid-like prophages, which are scattered throughout the chromosome.<sup>7</sup> The mosaic nature of the pathogenic *E. coli* genome suggests that virulence traits are acquired through the capture of mobile elements such as phages, insertion sequences, and plasmids. In particular, lambda-doid phage appears to have a major role in the emergence of EHEC strains, given that a variety of effector genes are encoded by several prophages lysogenized at different loci on the chromosome.<sup>7</sup> Therefore, for pathogenicity to be established, the expression kinetics of genes in different chromosomal loci must be appropriately orchestrated.

The transcription of LEE genes is regulated by two positive transcriptional regulators: Pch regulators activate the transcription of the *LEE1* operon, which encodes the transcriptional regulator Ler, which in turn regulates other LEE operons and genes in other loci.<sup>8</sup> Ler belongs to the H-NS family of nucleoid-associated proteins, exhibiting homology with the carboxy-termini of these proteins corresponding to DNA-binding domain.<sup>9</sup> Ler has been shown to activate LEE gene expression by counteracting the repression by H-NS.<sup>10</sup> EHEC strains possess five genes that encode homologs of Pch, though only two, *pchA* and *pchB*, activate *LEE1* transcription.<sup>8</sup> Pch is also a small protein as Ler, but it showed no homology with nucleoid-associated protein or transcription regulators. The expression of LEE genes is regulated in response to changes in environmental factors, such as the growth phase, temperature, osmolarity, bicarbonate ion concentration, membrane stress, and nutrient starvation.<sup>11–14</sup> Both Pch and Ler are major regulators of LEE gene expression in response to such changes. For example, transient growth arrest caused by insufficient nutrients or by entry into the stationary growth phase induces the expression of LEE genes through increased production of both Pch and Ler, because the transcription of *pch* genes and the *LEE1* operon is positively regulated by starvation-induced signal molecule, ppGpp.<sup>13</sup> Furthermore, the activation

of the RcsC–RcsD–RcsB phospho-relay system, which is thought to be a signal transduction system for responding membrane stress, affects the expression of LEE genes both positively and negatively through the activation of *grvA*, encoding a transcription regulator and the repression of *pch*.<sup>12</sup> Thus, environmental factors are evaluated by a sensing system common to all *E. coli* strains. In EHEC O157:H7, however, the signals from the sensing system also affect the transcription of O157-specific transcription regulators, which results in the coordinated regulation of the virulence genes in response to certain environmental changes.

In this study, in order to clarify which genes are involved in the pathogenicity of EHEC and the regulatory networks that govern the LEE and related genes, we isolated genes under the control of the two major virulence regulators, Pch and Ler, by comparing the changes in global transcription caused by altering the *pch* and *ler* expression levels. We identified two sets of genes whose transcription profiles were closely correlated with those of LEE genes. Furthermore, we determined the binding sites for PchA and Ler, and found that, although Pch was involved in the regulation of genes in both classes, Ler regulated the expression of only one class of genes.

## 2. Materials and methods

### 2.1. Bacterial strains and growth conditions

EHEC O157 Sakai (RIMD 0509952)<sup>3</sup> and its derivative strains used in this study are listed in Supplementary Table S1. Deletion mutants of EHEC O157 Sakai were constructed using the method and plasmids of Datsenko and Wanner.<sup>15</sup> DNA fragments corresponding to the *pchA* and *ler* genes were isolated by polymerase chain reaction (PCR) from the chromosomal DNA of O157 Sakai using specific primers: 5'-CACAGGAATATATCCGTACCC-3' and 5'-AGTATGTGTCACCTGGCCTATACGG-3' for *pchA*, and 5'-TTGGCTCACAATACTCATCC-3' and 5'-GCTTAACTAAATGGAAATGC-3' for *ler*. The PCR amplicon was cloned into pGEM-T (Promega) in the direction opposite to that of the *lac* promoter. The *pchA* gene of EHEC Sakai  $\Delta pchB \Delta pchC$  was tagged at the 3'-terminus with Strep-tag (IBA), and the *ler* gene in wild-type O157 Sakai was tagged with 3 × FLAG sequence, using an epitope-tagging procedure.<sup>16</sup> Bacteria were grown overnight in Luria-Bertani broth (LB), diluted 100-fold, and then cultured in LB or Dulbecco's modified Eagle's medium (DMEM) without fetal calf serum, at 37°C with shaking.

### 2.2. Transcriptome analysis

Total RNA from EHEC was isolated with TRI reagent (Sigma) using the method recommended by the

manufacturer. The RNA was purified again after treatment with RNase-free DNase I (Takara), and the amount and purity of the RNA were determined by measuring  $A_{260}$  and  $A_{280}$ . The RNAs from EHEC were used to synthesize cDNAs with Cy-3- or Cy-5-conjugated dUTP. cDNA labeling, microarray hybridization, data capture, and data analysis were performed as described previously.<sup>12</sup> We used a DNA microarray with oligo-DNAs specific for each of the CDS of EHEC O157 Sakai.<sup>12,17</sup> The expression ratio used here is to indicate the average of the ratios obtained in the two independent experiments in which the signal intensities were greater than the mean  $\pm 1$  SD of the negative control. The hierarchical clustering of gene expression profiles and conditions were analyzed using GeneSpring GX7.3 (Agilent Technologies). The raw and processed data are available in the GEO of NCBI (in accession no. GSE9752).

### 2.3. Design of EHEC O157 tiled microarrays

The EHEC O157 Sakai genome sequence and other genetic information were retrieved from the O157 Web site at NAIST (<http://genome.naist.jp/bacteria/o157/>) and used to design probes, following the Affymetrix guidelines (<http://www.affymetrix.com>). Two versions of the microarray were used in this study. One contained 25-nt-long probe sequences separated by 103 ( $\pm 5$ )-base intervals in the coding region and by intervals of 35 ( $\pm 2$ ) bases in the intergenic region along both strands of the O157 Sakai chromosome. The other contained 25-nt-long probe sequences separated by 45-base intervals, as possible, from only a single strand of the chromosome. If the sequence was not correct, the probe was offset with sequence within 10 bases. Thus, perfect-match probes and corresponding mismatch probes were synthesized on the chip.

### 2.4. ChIP-on-chip

Chromosome immunoprecipitation and DNA microarray analysis were performed as described by Oshima et al.<sup>18</sup> with slight modifications. EHEC O157 Sakai strain SK11172 or SK11282 was grown in DMEM to logarithmic phase ( $OD_{600} = 0.7 - 0.8$ ) at 37°C. Formaldehyde (final concentration of 1%) was added to 10 mL of the culture, and the mixture was incubated at room temperature for 20 min. To terminate the crosslinking reaction, glycine (final concentration of 0.45 M) was added, and the mixture was incubated at room temperature for 5 min. Bacteria were collected by centrifugation and washed first with Tris buffered saline, then with washing buffer [50 mM sodium phosphate (pH 8.0), 300 mM NaCl, 0.01% Tween-20] twice, suspended in 1 mL of lysis buffer [50 mM sodium phosphate (pH 8.0),

300 mM NaCl, 1% Tween-20, 20 mg/mL lysozyme], and stored at  $-20^{\circ}\text{C}$  until use. Bacteria were lysed by incubation at 37°C for 30 min and diluted in 4 mL of washing buffer containing phenylmethane-sulphonyl fluoride (final concentration of 1 mg/mL). The solution was then sonicated 15 times for 30 s at 30 s intervals at 4°C using a Bioruptor UCD-200 (Cosmo Bio Co., Ltd.). Cell debris was removed by centrifugation at 15 000 rpm for 30 min. An 800  $\mu\text{L}$  aliquot of the supernatant was mixed with magnetic beads bound to an anti-FLAG antibody (M2 monoclonal antibody, Sigma-Aldrich) or MagStrep (IBA) antibody and incubated at 4°C overnight. The beads were washed three times with washing buffer and then with TE (10 mM Tris-HCl (pH8.0), 1 mM EDTA). The protein-DNA complex bound to the beads was released in 100 mL of elution buffer [50 mM Tris-HCl (pH 7.5), 10 mM EDTA, 1% SDS] by heating at 65°C for 20 min for Ler-FLAG, or in Biotin buffer [50 mM Tris (pH 7.8), 150 mM NaCl, 5 mM EDTA, 10 mM Biotin] incubated at room temperature for 10 min for PchA-Strep. DNA fragments crosslinked to proteins were released by incubating the eluate with proteinase K (1 mg/mL) at 42°C for 2 h, then 65°C for 6 h, and purified using the Qiagen PCR Clean-up kit (Qiagen) in 200  $\mu\text{L}$  of TE. Recovered DNA fragments were amplified, terminally labeled, and hybridized with the EHEC O157 tiled microarray as described previously.<sup>18</sup> The raw intensity data (CEL files) were processed using the In Silico Molecular Cloning Array Edition software (In Silico Biology) for analysis and visualization. Signal intensities were obtained by subtracting the signal intensities of the mismatch probes from those of the perfect-match probes, and probes with negative values were excluded from further analysis. To compensate for differences among the probes caused by different probabilities of DNA disruption and hybridization efficiency, signals were normalized by dividing them by the signal intensities from DNA samples taken before the beads were added. The binding sites for PchA-Strep and Ler-FLAG were assessed manually. The reproducibility and specificity of binding of PchA or Ler to several novel sites were confirmed by ChIP-PCR for each site.

## 3. Results and discussion

### 3.1. Identification of Pch- and/or Ler-regulated genes

The genes in the LEE pathogenicity island are organized mainly into five operons, *LEE1* to *LEE5*, and several cistrons. The transcription of the LEE genes is positively regulated by the Ler protein, which is encoded by the first gene of the *LEE1* operon. Transcription of the *LEE1* operon is regulated by

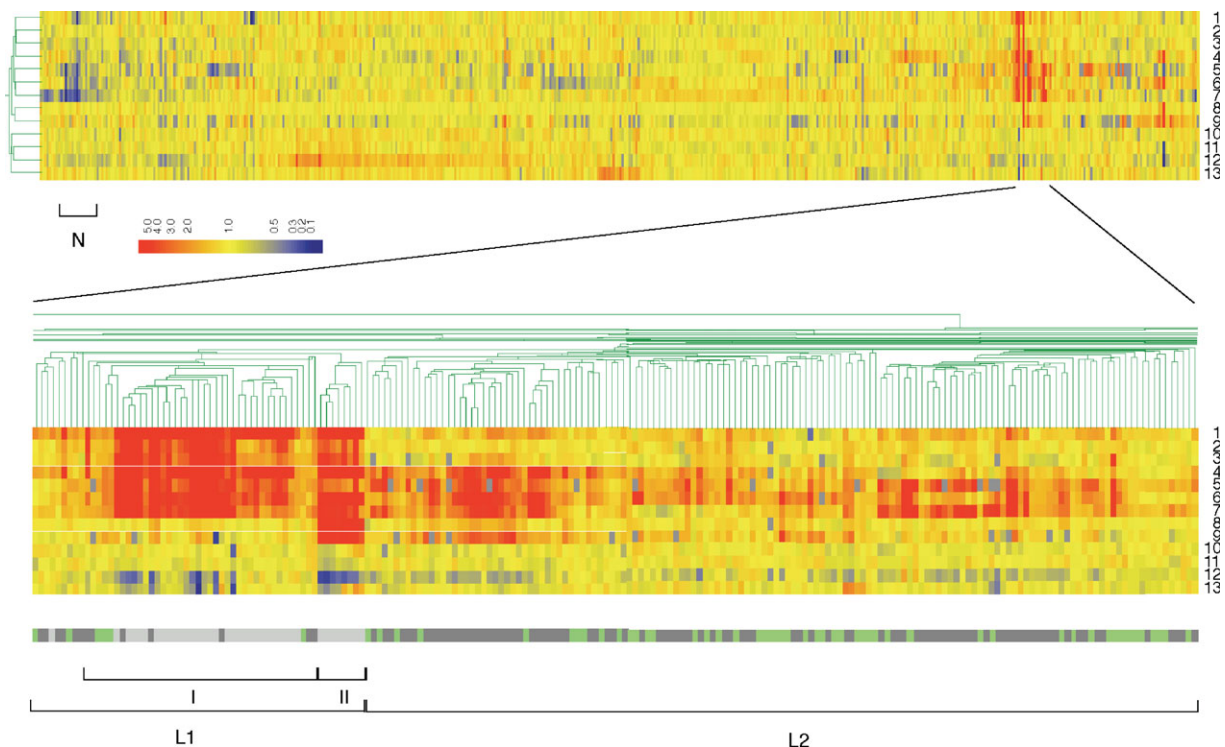


several proteins conserved in the non-pathogenic *E. coli* K12 strain, including IHF, QseA, and H-NS, and by the O157-specific regulator, Pch.<sup>8,14,19,20</sup> EHEC O157:H7 Sakai strain possesses three intact *pch* genes, which encode proteins that differ by only one or two amino acids, and are located in different prophage-like elements.

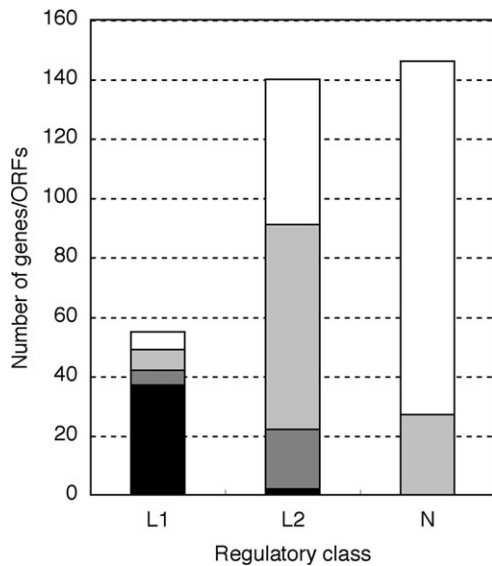
To elucidate the network created by O157-specific regulators, the effect of the overproduction or deletion of Ler or Pch on global transcription levels was determined using DNA microarray hybridization. We determined differences in the transcript levels for all the genes in EHEC O157 Sakai strain in 13 comparisons (Fig. 1): *ler* null mutant against wild-type EHEC, a *pchA-pchB-pchC* triple mutant against wild type, EHEC harboring a high copy number of the *ler* or *pchA* gene against EHEC harboring a control vector, the *ler* mutant harboring a high copy number of the *pchA* gene against the *ler* mutant harboring the control vector, the *pchA-pchB-pchC* triple mutant

harboring a high copy number of the *ler* gene against the *pch* triple-mutant harboring the control vector, and the *pchA-pchB* double mutant harboring a high copy number of *ler* against the double mutant harboring the control vector. Since the expression level of LEE genes in EHEC O157, when grown in tissue culture medium DMEM, is much higher than that in EHEC grown in LB, we included the differences in broth media as other factors affecting expression of EHEC genes. Except for the last case (grown in LB), all comparisons were carried out for bacteria grown in LB and DMEM.

The EHEC genes were subjected to clustering analysis by standard correlation to determine their transcript expression patterns (Fig. 1). As expected, genes belonging to the LEE showed similar expression profiles, and we classified them as part of a single group, class L1, which also included genes that were not encoded within the LEE (Fig. 2). Within the LEE-encoded genes of class L1, there were two distinct



**Figure 1.** Hierarchical cluster analysis of gene expression changes in response to changes in Pch or Ler levels. Gene expression data from 13 microarray experiments (lines) and whole EHEC O157 Sakai 5454 genes (columns) are indicated. The microarray experiments involved 13 comparisons between EHEC O157 Sakai and derivative strains: (1) Wild type (WT) harboring pGEM-*ler* to WT harboring pGEM vector grown in LB. (2) The *pchA-pchB* mutant harboring pGEM-*ler* to the *pchA-pchB* mutant harboring pGEM-vector grown in LB. (3) The *pchA-pchB-pchC* mutant harboring pGEM-*ler* to the *pchA-pchB-pchC* mutant harboring pGEM-vector grown in LB. (4) WT harboring pGEM-*pchA* to WT harboring pGEM-vector grown in LB. (5) WT harboring pGEM-*ler* to WT harboring pGEM-vector grown in DMEM. (6) WT harboring pGEM-*pchA* to WT harboring pGEM-vector grown in DMEM. (7) The *pchA-pchB-pchC* mutant harboring pGEM-*ler* to the *pchA-pchB-pchC* mutant harboring pGEM-vector grown in DMEM. (8) The *ler* mutant harboring pGEM-*pchA* to the *ler* mutant harboring pGEM vector grown in LB. (9) The *ler* mutant harboring pGEM-*pchA* to the *ler* mutant harboring pGEM vector grown in DMEM. (10) The *ler* mutant to WT grown in LB. (11) The *pchA-pchB-pchC* mutant to WT grown in LB. (12) The *pchA-pchB-pchC* mutant to WT grown in DMEM. (13) The *ler* mutant to WT grown in DMEM. The lower panel shows gene clusters in and around the LEE genes along with a color bar indicating the gene context (green: common to K12; dark grey: EHEC O157 specific; light grey: LEE). The scale bar indicates color coding of the RNA levels. Class L1 with subclasses I and II, class L2, and class N are indicated (see text).



**Figure 2.** Characteristics of genes in each cluster. Number of genes in three gene classes, L1, L2 and N, which were isolated by transcriptomic analysis, was shown. Genes were sorted by chromosomal locus and function as follows: black, LEE genes; grey, non-LEE virulence-associated genes; light grey, other genes in S-loop; white, genes on K12-common loci.

subgroups. One subgroup contained all the *LEE1* operon genes except *ler*, and the other contained most of the other LEE genes (Supplementary Table S2). We found that the transcript levels of the *LEE1* genes were increased by PchA overexpression in both wild-type EHEC and the *ler* deletion mutant, but the transcript levels of the other LEE genes were enhanced by PchA overexpression only in the presence of the *ler* gene. Because the transcript data were obtained from *ler* null mutant or from strains harboring multiple copies of *ler*, the *ler* gene itself could not be classified in this experiment.

The three class L1 effector genes that lie outside the LEE were *nleA*, *espF-U/tccP*, and *espJ*. The *nleA/espJ* gene is a single cistron on prophage Sp9 and is required for *Citrobacter rodentium*, an A/E mouse pathogen, to colonize the mouse colon.<sup>21</sup> *EspF-U/TccP* is necessary for the recruitment of actin to Tir beneath attached bacteria, which is essential for A/E lesion formation.<sup>22,23</sup> Two genes, *espF-U/tccP* and *espJ*, are located on lambdaoid-like prophage Sp14 and constitute an operon.<sup>24,25</sup> Since the *espF-M* gene of EHEC O157 Sakai is a pseudogene,<sup>23</sup> it is likely that *espF-M* was detected because of cross-hybridization with *espF-U/tccP*. Finally, class L1 included three unknown O157-specific genes in an S-loop, ECs0355, ECs0356, ECs0814, and the plasmid-born gene *tagA*. In addition, the transcription levels of three genes common to EHEC O157:H7 and K-12 in the *E. coli* backbone chromosome were regulated in close correlation with the LEE genes. These genes,

ECs0512 (*ylaD*), ECs0513 (*hha*), and ECs0514 (*ybaJ*), are likely to comprise an operon. The *hha* gene encodes a transcription regulator for *ler*.<sup>26</sup>

The expression profile of the class L1 cluster was correlated with that of another cluster of genes. We designated these genes as class L2 (Fig. 1). Class L2 contained 140 CDS/genes, consisting of 89 genes in an S-loop and 51 genes in the backbone or K-12 common chromosome (Fig. 2 and Supplementary Table S3). Although the transcript levels of the class L2 genes increased with increased Pch levels, like the genes in class L1, their response to increased levels of Ler was different, in that the L2 transcript levels increased only slightly with the overexpression of Ler when grown in LB. In contrast, the genes in L1 showed a marked increase in transcript level with high levels of Ler, even in LB.

Class L2 contained 20 non-LEE effector genes that encode T3SS-secreted proteins. Most of these effector genes are located on exchangeable effector loci (EEL) of lambda-like prophages,<sup>7</sup> including Sp3, Sp6, Sp9, Sp10, Sp11, Sp17, and SpLE3 (Fig. 2). This observation of the expression and T3SS-dependent secretion of these effectors when the EHEC harbored a high copy number of the *pchA* gene is consistent with our previous report.<sup>7</sup> All the effectors, except EspM2 (encoded by ECs3485), that were found in the culture supernatant of EHEC were regulated coordinately with the LEE genes. The expression of another effector gene, ECs4653, which encodes EspY4, was also upregulated by Ler and Pch, just like the other non-LEE effector genes.

Some genes were regulated in conjunction with their neighbors because of members of the same operon. One cluster of these genes, ECs0240–ECs0245, lies in the *rhsI* element, which is one of nine *rhs* elements in EHEC O157 Sakai. ECs0815 and ECs0816 seem to belong to the same operon and are transcribed in the opposite direction from ECs0814, which belongs to class L1. Another cluster of genes, from ECs3512 to ECs3508, seems to comprise an operon that lies at the end of prophage-like element Sp17. In addition, class L2 included genes for type-II secretion machinery (*etpC*, *etpD*, *etpE*, *etpH*, and *etpK*), *toxB*, and hemolysin (*hlyD*) on the virulence plasmid pO157 (Supplementary Table S3).

One cluster of genes showed a negative correlation in their transcription with changes in Ler or Pch expression (shown as class N in Fig. 1 and Supplementary Table S4). Most drastically affected were genes associated with an acid-resistant phenotype, which are closely clustered on the chromosome. These genes include *hdeA*, *hdeB*, *hdeD*, *gadE* (*yhiE*), *gadX*, and *gadW*. One hundred and nineteen genes conserved between K12 and EHEC O157 and 27 O157-specific genes on S-loops formed clusters

(Fig. 2). These genes showed very similar changes in transcript levels in response to changes in Ler/Pch levels. Several genes in this class, in addition to *gadE*, *gadX*, and *gadW*, regulate glutamate-dependent acid resistance. The EvgS–EvgA two-component regulatory system controls the expression of *gadE*, and transcription of the *gadXW* operon is dependent on alternative sigma factor RpoS, whose gene was also included in this class. Several of these factors, including GadE (YhiE), EvgS–EvgA, and RpoS, negatively affect the expression of LEE genes or bacterial adherence.<sup>27–29</sup> Although the ClpXP protease has been shown to contribute indirectly to the expression of the LEE genes,<sup>28</sup> the transcription of the *clpX–clpP* genes was negatively regulated by Ler and Pch.

### 3.2. Identification of binding sites for Pch and Ler

Although transcriptomic analyses revealed that Pch and/or Ler affect the transcription of many genes on S-loops and in the backbone or K-12 common chromosome, the transcript levels of some of them might be regulated indirectly, through proteins from genes that are directly regulated by Pch and/or Ler. To distinguish directly regulated genes from indirectly regulated genes, we determined the binding sites of PchA and Ler on the EHEC O157 Sakai chromosome by chromosome immunoprecipitation (ChIP) or chromosome affinity precipitation, in combination with DNA microarray hybridization. A derivative strain of EHEC O157 Sakai that expresses Strep-tagged PchA (PchA-Strep) or FLAG-tagged Ler (Ler-FLAG) was grown in DMEM to the logarithmic growth phase at 37°C, which is the optimal condition for expression of the LEE genes, and treated with formalin to crosslink the DNA with bound proteins. DNA fragments that bound to PchA-Strep or Ler-FLAG were separated from the bacterial lysate with Strep-Tactin-coated magnetic beads (MagStrep) or anti-FLAG-conjugated magnetic beads. After decrosslinking and purification, the DNA fragments that bound to PchA or Ler were mapped by hybridization with a high-density tiled microarray. We found 259 loci for Pch and 59 for Ler on the chromosome (Fig. 3 and Supplementary Tables S5 and S6). Of these, 43.1% of the Pch-bound loci and 39.0% of the Ler-bound loci were in S-loops or junctions with S-loops and in the backbone or K12-common chromosome. Considering that only 25.5% of the chromosome was unique to the EHEC O157 Sakai strain compared with *E. coli* K12 strain MG1655, it was clear that Pch and Ler preferentially bound to S-loop sequences rather than to the backbone or K12-common chromosome.

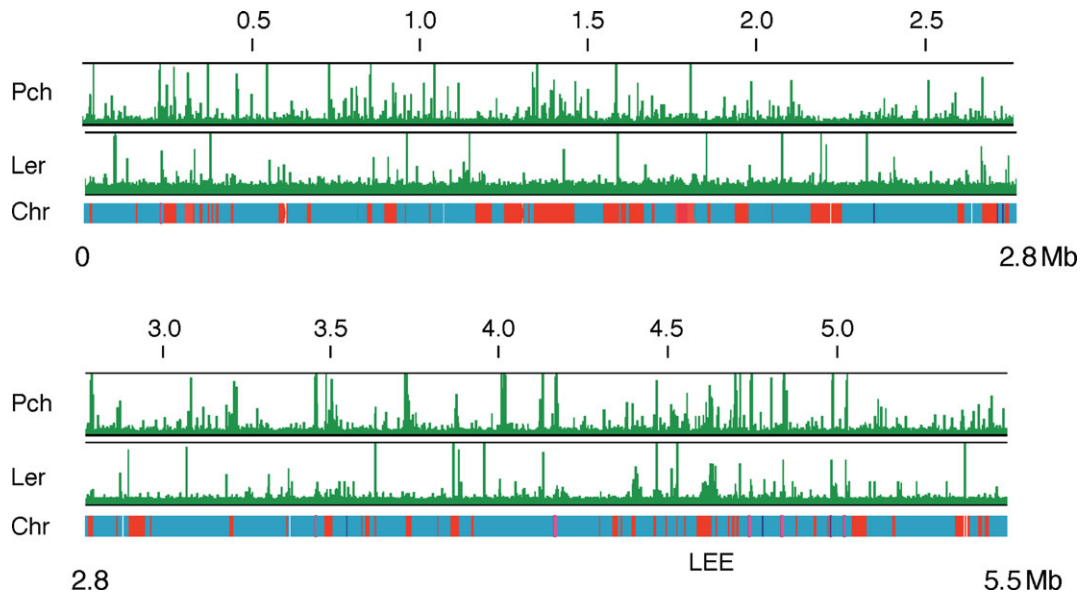
The expression of the LEE genes is positively regulated by a molecular cascade: Pch is a positive

transcriptional regulator for the *LEE1* promoter.<sup>8</sup> Activation of the *LEE1* operon leads to the expression of Ler, which activates the other LEE genes in operons *LEE2* to *LEE5* and several cistrons. As expected, Pch bound to the *LEE1* promoter region (Fig. 4A). In addition, Pch bound to sites throughout the LEE sequence, including the coding region of the *LEE1* operon. Furthermore, although the binding was weaker than in the *LEE1* operon, Pch bound to other parts of the LEE island: the upstream region of the *grlR–grlA* operon, the 5' portions of the *LEE2* and *LEE3* operons, the downstream portion of the *LEE3* operon, the upstream portion of *tir*, the coding region of *cesT*, and the downstream portion of *eaE*. Ler bound at similar levels to several regions in the LEE island, mostly in the 5' regions of operons or genes (Fig. 4A): the intergenic regions between *LEE1* and *espG* and between *grlR* and ECs4579 (*rorf3*), the 5' portions of the *LEE2* and *LEE3* operons, the intergenic region between *cesF* and *map*, and the 5' regions of *escD* and *sepL*.

EHEC O157 Sakai possesses 24 prophage or prophage-like elements, which make up more than half of the S-loop sequences. The binding loci for Pch and Ler were unevenly distributed in these acquired sequences. They rarely bound to sequences conserved among the phages; rather, they mostly bound to sequences that were unique to each phage element. Nine lambdaoid-like phage elements and two phage-like elements possess EEL, which encode the majority of functional effector genes. Binding loci of Pch were detected within the EEL of all phage elements and Ler-binding loci were detected within the EEL of Sp9, Sp14, and SpLE4 (including LEE) (Fig. 4). Among EHEC O157-specific loci, the locus that contains genes for synthesis of polysaccharide moieties of lipopolysaccharide was bound by both Pch and Ler. Other locus that was covered by Pch was ETT2 locus encodes incomplete type III secretion system with truncated substrate genes. Fourteen loci encoding a set of fimbriae biosynthesis genes were identified on EHEC O157 Sakai chromosome. Pch was bound upstream part of 11 fimbriae operons and three of them were also bound by Ler. They include three loci that are conserved in K-12 strain and four loci that are partially conserved in K-12 strain.

Although most regions bound by Pch ranged from 0.3 to 0.8 kb, several loci showed binding that extended to as much as 3.0 kb. Of the 59 Ler-binding loci, 46 overlapped with the Pch-binding loci, and all the Ler-binding loci in S-loops also bound Pch. Because of the low resolution of ChIP-chip analysis, we could not identify a consensus sequence for the Pch-binding sites other than that they are of lower G + C content than average of *E. coli* chromosome. These results suggested that





**Figure 3.** Distribution of PchA and Ler binding over the EHEC O157 Sakai chromosome. Graphs show PchA and Ler binding to the chromosome of the EHEC O157 Sakai strain (horizontal bar). Red and blue segments indicate S-loops (O157:H7 strain-specific regions) and backbone (regions conserved with *E. coli* K12), respectively. Vertical bars indicate the relative hybridization intensity of the precipitated DNA compared with DNA in the supernatant.

both Pch and Ler are DNA-binding proteins with a low binding specificity.

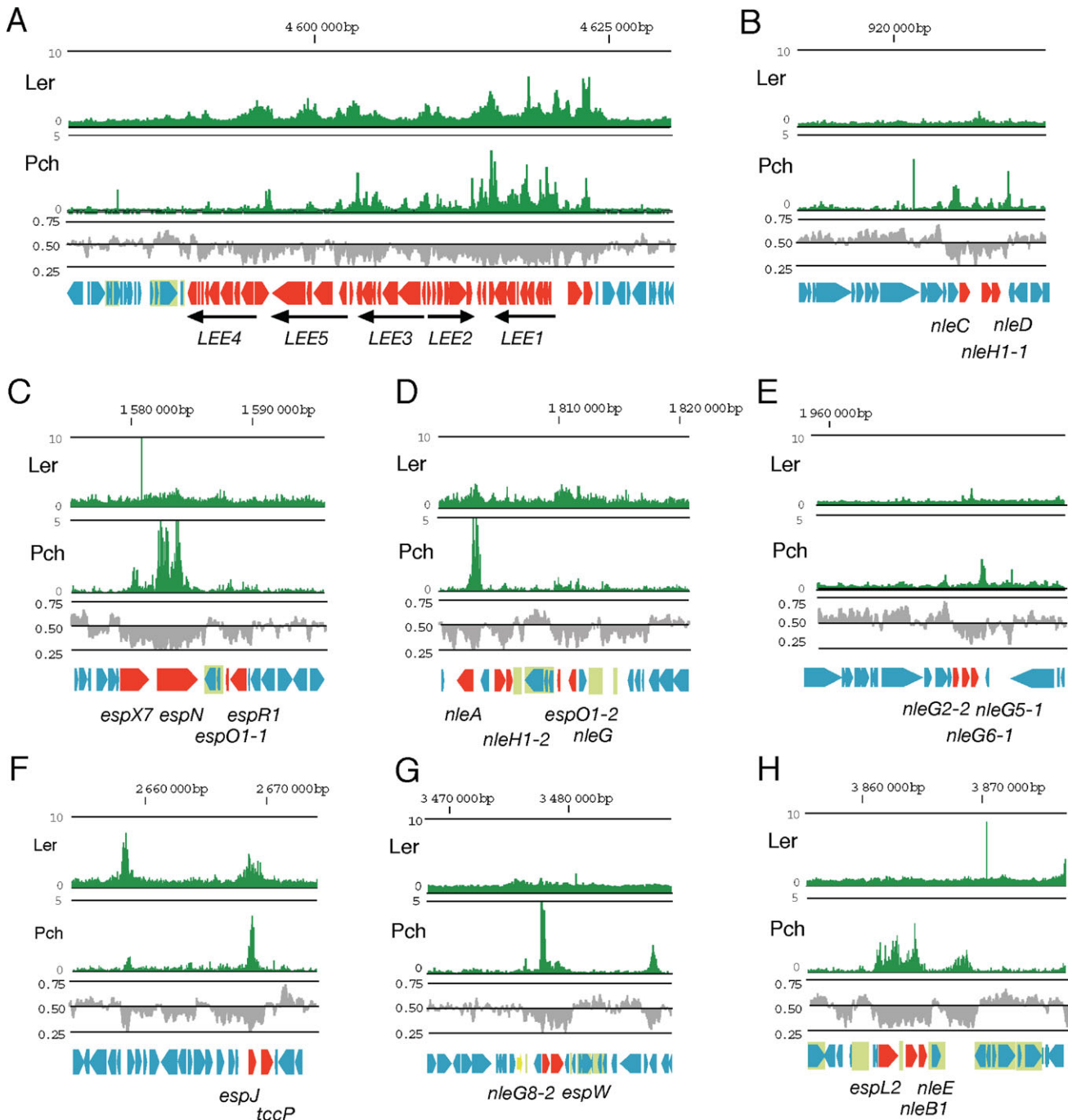
### 3.3. Correlation between gene transcription profiles and binding by Pch and/or Ler

We next considered whether there were differences in binding by Pch and Ler in the L1 and L2 gene classes. Although most of effector genes were transcriptionally activated by Pch, they were classified into two distinctive classes by transcription profiles. Ten effector genes in class L1—seven in the LEE island and three outside it—were regulated coordinately with other LEE genes that function in the biosynthesis of the type III secretion machinery. Both Pch and Ler bound to the upstream or coding sequences of all three L1 prophage-associated effector genes, *nleA* (ECs1812), *espJ* (ECs2714), and *espF-u/tccP* (ECs2715) (Fig. 4D and F), strongly suggesting the direct activation of their transcription through coordinate binding by Pch and Ler.

Twenty-one class L2 effector genes were located on seven prophage-like elements scattered throughout the chromosome. In contrast to class L1 effector genes, although Pch bound in and around the L2 effector gene sequences, Ler binding was not detected in many cases (Fig. 4B–E, G, and H). Positions of binding loci of Pch relative to effector genes vary among genes. For example, Pch bound upstream of *nleC*, *nleA*, *nleG8-2*, whereas binding to coding regions were observed for *espX7*, *espN*, *espJ*, and *espL2*. In the EEL of Sp10, two Pch-binding loci were

found upstream and downstream of a cluster of three *nleG* genes. Three L2 effector genes, *nleG5-2* (ECs2154), *nleG6-2* (ECs2155), and *nleG2-3* (ECs2156), on Sp11 bound Pch only at upstream of the gene cluster. Since the probes for *nleG5-2*, *nleG6-2*, and *nleG2-3* used in the DNA microarray for transcriptome analysis were the same as for *nleG5-1* (ECs1996), *nleG6-1* (ECs1995), and *nleG2-2* (ECs1994) on Sp11, respectively, it is likely that the signals we detected originated only from the genes on Sp10 and that the genes on Sp11 were less expressed. These binding data suggested that the effector genes in class L2 were directly activated by Pch without Ler.

Of the 57 genes in class L1, only three S-loop genes and two K-12 common genes showed no binding by PchA or Ler (Supplementary Table S1). Rather, 68.6% of the L1 S-loop genes bound both Pch and Ler, as did 50% of the L1 backbone or K-12 common genes (Fig. 5). Therefore, most class L1 genes, including those on the backbone, are probably directly regulated by Pch and Ler. However, among the L2 genes, more than half bound neither Pch nor Ler, suggesting that their transcription was regulated indirectly by these proteins. As shown for the effector genes, the non-effector class L2 genes that bound Pch, but not Ler, were much more abundant than those that bound both (Fig. 5 and Table 2). 66.6% of Pch-bound genes were not bound by Ler. This suggested that the transcriptional regulation of the class L2 genes was, overall, directly regulated by Pch without the involvement of Ler.

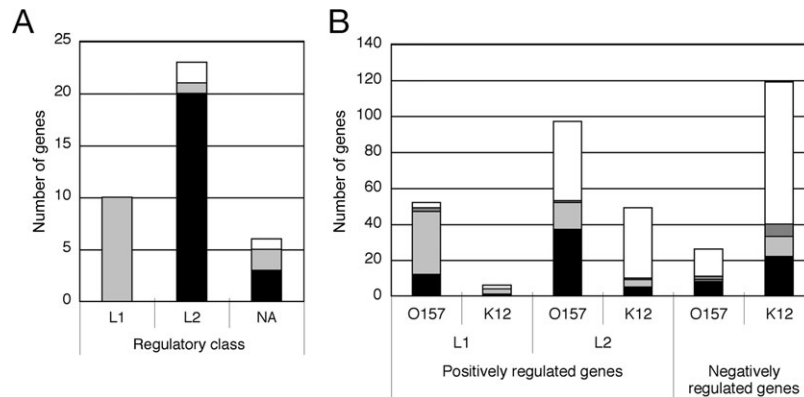


**Figure 4.** Distribution of PchA and Ler binding in the LEE and EELs. Graphs show PchA and Ler binding as in Fig. 2. Upper line indicates the position from the replication origin of the chromosome. Middle row indicates G + C contents. **(A)** Chromosomal region around the LEE. Genes in the LEE are represented by red arrows. The five main LEE operons, *LEE1* to *LEES*, are represented by arrows below the open reading frames map. **(B)** EEL on prophage-like element Sp3. Genes encoding effector proteins are indicated by red arrows. **(C)** EEL on prophage-like element Sp6. **(D)** EEL on prophage-like element Sp9. **(E)** EEL on prophage-like element Sp10. **(F)** EEL on prophage-like element Sp14. **(G)** EEL on prophage-like element Sp17. **(H)** EEL on prophage-like element SpLE3. IS and remnant of IS are shown in light green boxes.

Although 146 genes were classified as class N, that is, their expression was negatively correlated with Pch and Ler activity, only 35% showed Pch and/or Ler binding in or near their DNA sequences (Fig. 5). These 40 backbone or K-12 common genes and 11

S-loop genes seemed to be directly regulated by Pch and/or Ler (Table 3). They include genes encoding regulators, *yhiF*, *gadE*, *gadX*, and *gadW*, and their regulating genes, *gadB*, *hdeA*, *hdeB*, and *hdeD*. Two genes, *yhiF* and *gadE*, inhibit the expression of LEE genes,<sup>27</sup>





**Figure 5.** Correlation between Pch and/or Ler binding with transcription profiles. The binding of Pch and Ler to loci in and around genes was determined as in Fig. 2 and Fig. 3. **(A)** The number of effector genes assigned to class L1 and class L2. Effector genes not in either class are labeled NA (not applicable). **(B)** The number of genes in classes L1, L2, and N. Genes in S-loops (EHEC O157 strain-specific regions) and in the *E. coli* backbone (regions conserved with *E. coli* K-12) are identified as O157 and K12, respectively. Binding of Pch and/or Ler is denoted as follows: grey, bound by Ler and Pch; black, bound by Pch, but not Ler; dark grey, bound by Ler, but not Pch; white, no binding by Pch or Ler.

suggesting there is a reciprocal negative interaction between the Pch–Ler virulence regulatory system and the YhiF–GadE regulatory system.

#### 3.4. Global regulation to promote the efficient expression of virulence traits

Pch is a 12-kDa protein without significant homology to any transcriptional regulatory proteins except PerC (also called BfpW)<sup>30,31</sup> which is encoded on a virulence plasmid of enteropathogenic *E. coli*. The *pch* genes encode a transcriptional activator for the *LEE1* promoter, and the deletion of *pchA* and *pchB* genes greatly reduces the production and secretion of EspB and EspA.<sup>8</sup> The introduction of multiple copies of the *pchA* gene into EHEC O157 enhances the production and secretion of effector proteins, including many non-LEE encoded effectors,<sup>7</sup> indicating that Pch is involved in the regulation of non-LEE effector genes as well as of LEE genes. Our study clearly demonstrated that Pch acts as a global regulator of many genes scattered throughout the EHEC O157 chromosome and plasmids. As expected, most effector genes in the EEL on prophage-like elements were positively regulated by Pch. The molecular mechanism of transcriptional activation by Pch may not be simple, since the position of its binding site varied from gene to gene, as did the number of binding sites. Therefore, Pch probably uses a transcriptional activating mechanism that is different from that of a typical transcriptional activator, like an AraC-like protein, which activates transcription by binding upstream of the promoter and interacting with RNA polymerase. The binding of Pch in the coding region or downstream of its target genes suggested that Pch affects promoter activity

without directly interacting with RNA polymerase, most likely by inducing changes in the chromosomal structure around the promoter. We could not identify a consensus sequence for Pch binding, besides low G + C content, from the ChIP on chip data, because of low resolution. To characterize more precisely, further analysis such as footprinting must be required.

The analysis of transcription profiles revealed that the transcriptional responses of genes in class L1 to changes in the levels of Pch and Ler were slightly different from those of the class L2 genes. Ler binding may be required to activate most class L1 promoters, and these promoters were activated in the absence of Pch as efficiently as in its presence. Although the transcript levels of several L2 genes could be increased in a Pch-negative background by overexpressing Ler, the increase was much smaller than that obtained by the overexpression of Pch. It is likely that Ler was primarily a regulator for a restricted group of genes that are functionally related to the A/E phenotype as compared to Pch. It is supported by one of the non-LEE-encoded effectors, TccP/EspFu, which is required for actin polymerization at the cytoplasmic domain of Tir.<sup>22,23</sup> In addition, the *nleA* (*espl*) gene of *C. rodentium*, a mouse A/E pathogen, is involved in the colonization of mouse colon and the induction of hyperplasia.<sup>21,32</sup> Finally, since the *ler* gene is LEE gene, the Ler target genes might primarily be LEE genes.

In the process of establishing the pathogenicity of EHEC, genes encoding proteins with A/E phenotype-related functions might be selected and placed under the direct control of Ler. Pch seems to have a broader role in regulating gene expression, since its binding sites were scattered over the chromosome. EHEC O157:H7 Sakai possesses five *pch* genes,

**Table 1.** ORFs in cluster L1 with binding sites for Pch/Ler

ECs#	Gene/ORF	Product	Locus	Bound by***
ECs0207	<i>dniR</i>	Transcriptional regulator for nitrite reductase (cytochrome c552)	K12	P
ECs0335		Putative oxidoreductase	S-loop57	P
ECs0354		Putative invertase	S-loop54	P + L
ECs0355		Putative invertase	S-loop54	P + L
ECs0356		Hypothetical protein	S-loop54	P + L
ECs0512	<i>ylaD</i>	Maltose O-acetyltransferase	K12	P + L****
ECs0513	<i>hha</i>	Hemolysin expression modulating protein	K12	P + L****
ECs0514	<i>ybaJ</i>	Hypothetical protein	K12	P + L
ECs0814		Putative outer membrane protein	Sp3	P + L
ECs1126/2715*	<i>espF-M/U</i>	Type-III secretion effector EspF-U/TccP	Sp4/Sp14	P + L****
ECs1812*	<i>nleA/espl</i>	Type-III secretion effector NleA/Espl	Sp9	P + L
ECs2714*	<i>espJ</i>	Type-III secretion effector EspJ	Sp14	P + L
ECs3500		Hypothetical membrane protein	Sp17	P
ECs4550*	<i>espF</i>	Effector protein EspF	LEE	L****
ECs4551	L0017	T3SS protein	LEE	L****
ECs4552	<i>escF</i>	T3SS EscF protein	LEE	L****
ECs4553	<i>cesD2</i>	T3SS chaperone CesD2	LEE	L****
ECs4554*	<i>espB</i>	Translocator protein EspB	LEE	L****
ECs4555*	<i>espD</i>	Translocator protein EspD	LEE	L****
ECs4556*	<i>espA</i>	Translocator protein EspA	LEE	L****
ECs4557	<i>sepL</i>	T3SS SepL protein	LEE	L
ECs4558	<i>escD</i>	T3SS EscD protein	LEE	P + L
ECs4559	<i>eae</i>	Gamma intimin	LEE	P + L****
ECs4560	<i>cesT</i>	T3SS chaperon CesT	LEE	P + L
ECs4562*	<i>map</i>	Effector protein Map	LEE	L
ECs4563	<i>cesF</i>	T3SS chaperon CesF	LEE	P + L****
ECs4564*	<i>espH</i>	Effector protein EspH	LEE	P + L****
ECs4565	<i>sepQ</i>	T3SS SepQ protein	LEE	P + L****
ECs4566	L0032	Hypothetical protein	LEE	P + L****
ECs4567	L0033	Hypothetical protein	LEE	P + L****
ECs4568	<i>escN</i>	T3SS protein EscN	LEE	P + L
ECs4570	L0036	Hypothetical protein	LEE	P + L
ECs4571*	<i>sepZ</i>	Effector protein SepZ/EspZ	LEE	P + L
ECs4572	L0038	Hypothetical protein	LEE	P + L****
ECs4573	<i>escJ</i>	T3SS EscJ protein	LEE	P + L****
ECs4574	<i>sepD</i>	T3SS protein SepD	LEE	P + L****
ECs4575	<i>escC</i>	T3SS EscC protein	LEE	P + L****
ECs4576	<i>cesD</i>	T3SS chaperone CesD	LEE	P + L****
ECs4577	<i>grlA</i>	Transcription regulator protein GrlA	LEE	P + L
ECs4578	<i>grlR</i>	Transcription regulator protein GrlR	LEE	P + L
ECs4579	L0045	Hypothetical protein	LEE	P + L
ECs4580	<i>escU</i>	T3SS EscU protein	LEE	P + L
ECs4581	<i>escT</i>	T3SS EscT protein	LEE	P + L
ECs4582	<i>escS</i>	T3SS EscS protein	LEE	P + L
ECs4583	<i>escR</i>	T3SS EscR protein	LEE	P + L

Continued

**Table 1.** Continued

ECs#	Gene/ORF	Product	Locus	Bound by***
ECs4584	L0050	Hypothetical protein	LEE	P + L
ECs4585	L0051	Hypothetical protein	LEE	P + L
ECs4586	L0052	Hypothetical protein	LEE	P + L
ECs4587	L0053	Hypothetical protein	LEE	P + L
ECs4590*	<i>espG</i>	Effector protein EspG	LEE	P + L

\*Gene encoding type-III effector or translocator protein.

\*\*ECs4588 (Ier) was not included because of artificially controlled expression.

\*\*\*P, Pch; L, Ler; P + L, Pch and Ler.

\*\*\*\*Bound at upstream gene in operon.

**Table 2.** ORFs in clusters L2 with binding sites for Pch/Ler

ECs#	Gene/ORF	Product	Locus	Binding**
ECs0240		Hypothetical membrane protein	S-loop13	P + L
ECs0241		H repeat-associated protein	S-loop13	P + L***
ECs0242		Rhs core protein	S-loop13	P + L ***
ECs0243		Hypothetical protein	S-loop13	P + L***
ECs0244		Hypothetical protein	S-loop13	P + L***
ECs0245		Putative H repeat-associated protein	S-loop13	L
ECs0293		Hypothetical protein	S-loop19	P + L
ECs0294		Hypothetical protein	S-loop19	P + L
ECs0295		Hypothetical protein	S-loop19	P + L
ECs0815		Antitermination protein	Sp3	P + L
ECs0816		Hypothetical membrane protein	Sp3	P + L***
ECs0846	<i>nleB2-1'</i>	NleB-like protein, pseudogene	Sp3	P
ECs0848*	<i>nleH1-1</i>	Effector protein <i>NleH1</i>	Sp3	P
ECs0849		Hypothetical protein	Sp3	P
ECs0949		Hypothetical protein	S-loop74	P
ECs1560*	<i>espX7</i>	Effector protein EspX7	Sp6	P
ECs1561*	<i>espN</i>	Effector protein EspN	Sp6	P
ECs1562		Hypothetical protein	Sp6	P***
ECs1567*	<i>espO1-1</i>	Effector protein EspO1-1	Sp6	P
ECs1568*	<i>espK</i>	Effector protein EspK	Sp6	P
ECs1814*	<i>nleH1-2</i>	Effector protein NleH1-2	Sp9	P***
ECs1815*	<i>nleF</i>	Effector protein NleF	Sp9	P
ECs1821	<i>espO1-2</i>	Effector protein <i>EspO1-2</i>	Sp9	P + L
ECs1822		Hypothetical protein	Sp9	P + L
ECs1823		Hypothetical protein	Sp9	P + L
ECs1824*	<i>nleG</i>	Effector protein NleG	Sp9	P + L***
ECs1994/ECs2156*	<i>nleG2-2/nleG2-3</i>	Effector protein NleG2-2/NleG2-3	Sp10/Sp11	P***
ECs1995/ECs2155*	<i>nleG6-1/nleG6-2</i>	Effector protein NleG6-1/NleG6-2	Sp10/Sp11	P***
ECs1996/ECs2154*	<i>nleG5-1/nleG5-2</i>	Effector protein NleG5-1/NleG5-2	Sp10/Sp11	P
ECs2701	<i>yedR</i>	Predicted inner membrane protein	K12	P + L
ECs2917	<i>yehD</i>	Putative fimbrial-like protein	K12	P + L

Continued



Table 2. Continued

ECs#	Gene/ORF	Product	Locus	Binding**
ECs3140	<i>ais</i>	Aluminum-inducible protein	K12	P
ECs3367	b2505	Putative outer membrane lipoprotein	K12	L
ECs3486*	<i>nleG8-2</i>	Effector protein NleG8-2	Sp17	P
ECs3487*	<i>espW</i>	Effector protein EspW	Sp17	P
ECs3508		Hypothetical protein	Sp17	P
ECs3509		Hypothetical protein	Sp17	P
ECs3510		Hypothetical protein	Sp17	P
ECs3511		Hypothetical protein	Sp17	P
ECs3512		Putative site-specific recombinase	Sp17	P
ECs3517	<i>ypjC</i>	Hypothetical protein	K12	P + L
ECs3616	<i>ybdY</i>	Hypothetical protein	K12	P + L
ECs3703	<i>yqeH</i>	Hypothetical protein	K12	P
ECs3713	b2857	Hypothetical protein	K12	P
ECs3718	<i>eprI</i>	Type III secretion protein EprI	ETT2	P
ECs3720		Putative transcriptional regulator	ETT2	P
ECs3721	<i>epaS</i>	Type III secretion protein EprS	ETT2	P
ECs3855*	<i>espL2</i>	Effector protein EspL2	SpLE3	P
ECs3857*	<i>nleB1</i>	Effector protein NleB1	SpLE3	P
ECs3858*	<i>nleE</i>	Effector protein NleE	SpLE3	P
ECs4000	<i>yhaB</i>	Hypothetical protein	K12	P
ECs4001	<i>yhaC</i>	Hypothetical protein	K12	P
ECs4292		Hypothetical membrane protein	S-loop87	P
ECs4569	<i>escV</i>	T3SS EscV protein	LEE	P
ECs4589		Hypothetical protein	LEE	P + L
ECs4653*	<i>espY4</i>	Effector protein EspY4	S-loop39	P
ECs4656		Hypothetical protein	S-loop39	P
ECs4657	<i>espY5'</i>	Pseudogene	S-loop39	P
ECs4658		Hypothetical protein	S-loop39	P
ECs5314		Hypothetical protein	S-loop	P

\*Gene encoding type-III effector protein (Tobe et al., 2006).

\*\*P, Pch; L, Ler; P + L; Pch and Ler.

\*\*\*Bound at upstream gene in operon.

including at least two encoding active proteins and two truncated pseudogenes, on different prophage-like elements.<sup>8,13</sup> Furthermore, multiple *pch* genes are found among EHEC O157 strains and other serotypes of EHEC strains, such as O26, O111, and O103 (Ogura et al., unpublished). These findings suggest that Pch is required for the expression of virulence genes and that the *pch* gene itself is rather unstable.

Horizontally transferred genes are thought to be initially repressed by the binding of H-NS to the DNA around them.<sup>18,33</sup> Pch might counteract the effect of H-NS, both by activating the horizontally acquired genes and influencing the expression of genes in the backbone or K-12 common chromosome. The global effect of Pch on both EHEC-specific

and K-12 common genes must be to promote the integration of the acquired genes with a virulence regulon and to govern the optimization of virulence expression. Because changing the expression levels of some K-12 common genes could have an adverse effect on bacterial physiology, the virulence-regulatory system must be strictly controlled and activated only when it is required. In fact, the evidence shows that the activation of the virulence-regulatory system is strictly regulated in response to environmental signals, through the expression of both *pch* and *ler*.<sup>12,13</sup> Thus, the virulence-regulatory system involving Pch and Ler is a good model for understanding the process of the evolution of pathogenicity and the emergence of novel pathogens.

**Table 3** ORFs in cluster N with binding sites for Pch/Ler

ECs#	Gene/ORF	Product	Locus	Binding*
ECs0079	<i>leuL</i>	Leu operon leader peptide	K12	P + L
ECs0324	<i>ykgK</i>	Putative regulator	K12	L
ECs0538	<i>ybaS</i>	Putative glutaminase	K12	P
ECs0539	<i>ybaT</i>	Putative amino acid/amine transport protein	K12	P
ECs1041	<i>ompA</i>	Outer membrane protein 3a	K12	P
ECs1272		Rtn-like protein	S-loop	P
ECs2603	<i>yecG</i>	Putative regulator	K12	P
ECs2712	<i>yodB</i>	Putative cytochrome	K12	P
ECs2846	<i>galF</i>	Putative Galf transferase	K12	P
ECs2974	<i>stx1A</i>	Shiga toxin I subunit A precursor	Sp15	P
ECs3104	<i>ompC</i>	Outer membrane protein C OmpC	K12	P + L
ECs3271	b2392	High affinity manganese transport protein	K12	P + L
ECs3272	<i>nupC</i>	Permease of transport system for 3 nucleosides	K12	P + L
ECs3284	<i>zipA</i>	Cell division protein involved in FtsZ ring	K12	P
ECs3289	<i>crr</i>	Glucose-specific PTS system IIA component	K12	P
ECs4045	<i>pnp</i>	Polynucleotide phosphorylase	K12	P
ECs4112	<i>yhcO</i>	Hypothetical protein	K12	L
ECs4377	<i>slp</i>	Huter membrane protein Slp	K12	P + L
ECs4379		Hypothetical protein	S-loop	P + L
ECs4380		Heme utilization/transport protein	S-loop	P + L
ECs4389	<i>hdeB</i>	Hypothetical protein	K12	P + L
ECs4390	<i>hdeA</i>	Hypothetical protein	K12	P + L
ECs4391	<i>hdeD</i>	Hypothetical protein	K12	P + L
ECs4392	<i>yhiE</i>	Hypothetical protein	K12	P + L
ECs4393	<i>yhiU</i>	Putative membrane protein	K12	P + L
ECs4395	<i>gadW</i>	Putative ARAC-type regulatory protein	K12	P + L
ECs4396	<i>gadX</i>	Putative ARAC-type regulatory protein	K12	P + L
ECs4427		Putative fimbrial protein precursor	S-loop	L

\*P, Pch; L, Ler; P + L, Pch and Ler.

**Supplementary data:** Supplementary data are available online at [www.dnaresearch.oxfordjournals.org](http://www.dnaresearch.oxfordjournals.org).

## Funding

This work was supported by Grants-in-Aid for Scientific Research on Priority Areas (C) from The Ministry of Education, Culture, Sports, Science, and Technology of Japan.

## References

- Gal-Mor, O. and Finlay, B. B. 2006, Pathogenicity islands: a molecular toolbox for bacterial virulence, *Cell Microbiol.*, **8**, 1707–1719.
- Kaper, J. B., Nataro, J. P. and Mobley, H. L. 2004, Pathogenic *Escherichia coli*, *Nat. Rev. Microbiol.*, **2**, 123–140.
- Hayashi, T., Makino, K., Ohnishi, M., et al. 2001, Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12, *DNA Res.*, **8**, 11–22.
- Frankel, G., Phillips, A. D., Rosenshine, I., et al. 1998, Enteropathogenic and enterohaemorrhagic *Escherichia coli*: more subversive elements, *Mol. Microbiol.*, **30**, 911–921.
- Hayward, R. D., Leong, J. M., Koronakis, V., et al. 2006, Exploiting pathogenic *Escherichia coli* to model transmembrane receptor signalling, *Nat. Rev. Microbiol.*, **4**, 358–370.
- Deng, W., Puente, J. L., Gruenheid, S., et al. 2004, Dissecting virulence: systematic and functional analyses of a pathogenicity island, *Proc. Natl. Acad. Sci. USA*, **101**, 3597–3602.

7. Tobe, T., Beatson, S. A., Taniguchi, H., et al. 2006, An extensive repertoire of type III secretion effectors in *Escherichia coli* O157 and the role of lambdoid phages in their dissemination, *Proc. Natl. Acad. Sci. USA*, **103**, 14941–14946.
8. Iyoda, S. and Watanabe, H. 2004, Positive effects of multiple pch genes on expression of the locus of enterocyte effacement genes and adherence of enterohaemorrhagic *Escherichia coli* O157:H7 to HEp-2 cells, *Microbiology*, **150**, 2357–2571.
9. Elliott, S. J., Sperandio, V., Giron, J. A., et al. 2000, The locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE- and non-LEE-encoded virulence factors in enteropathogenic and enterohemorrhagic *Escherichia coli*, *Infect. Immun.*, **68**, 6115–6126.
10. Bustamante, V. H., Santana, F. J., Calva, E., et al. 2001, Transcriptional regulation of type III secretion genes in enteropathogenic *Escherichia coli*: Ler antagonizes H-NS-dependent repression, *Mol. Microbiol.*, **39**, 664–678.
11. Abe, H., Tatsuno, I., Tobe, T., et al. 2002, Bicarbonate ion stimulates the expression of locus of enterocyte effacement-encoded genes in enterohemorrhagic *Escherichia coli* O157:H7, *Infect. Immun.*, **70**, 3500–3509.
12. Tobe, T., Ando, H., Ishikawa, H., et al. 2005, Dual regulatory pathways integrating the RcsC–RcsD–RcsB signalling system control enterohaemorrhagic *Escherichia coli* pathogenicity, *Mol. Microbiol.*, **58**, 320–333.
13. Nakanishi, N., Abe, H., Ogura, Y., et al. 2006, ppGpp with DksA controls gene expression in the locus of enterocyte effacement (LEE) pathogenicity island of enterohaemorrhagic *Escherichia coli* through activation of two virulence regulatory genes, *Mol. Microbiol.*, **61**, 194–205.
14. Umanski, T., Rosenshine, I. and Friedberg, D. 2002, Thermoregulated expression of virulence genes in enteropathogenic *Escherichia coli*, *Microbiology*, **148**, 2735–2744.
15. Datsenko, K. A. and Wanner, B. L. 2000, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products, *Proc. Natl. Acad. Sci. USA*, **97**, 6640–6645.
16. Uzzau, S., Figueroa-Bossi, N., Rubino, S., et al. 2001, Epitope tagging of chromosomal genes in *Salmonella*, *Proc. Natl. Acad. Sci. USA*, **98**, 15264–15269.
17. Ogura, Y., Kurokawa, K., Ooka, T., et al. 2006, Complexity of the genomic diversity in enterohemorrhagic *Escherichia coli* O157 revealed by the combinational use of the O157 Sakai Oligo-DNA microarray and the whole genome PCR scanning, *DNA Res.*, **13**, 3–14.
18. Oshima, T., Ishikawa, S., Kurokawa, K., et al. 2006, *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase, *DNA Res.*, **13**, 141–153.
19. Friedberg, D., Umanski, T., Fang, Y., et al. 1999, Hierarchy in the expression of the locus of enterocyte effacement genes of enteropathogenic *Escherichia coli*, *Mol. Microbiol.*, **34**, 941–952.
20. Sharp, F. C. and Sperandio, V. 2007, QseA directly activates transcription of LEE1 in enterohemorrhagic *Escherichia coli*, *Infect. Immun.*, **75**, 2432–2440.
21. Mundy, R., Petrovska, L., Smollett, K., et al. 2004, Identification of a novel *Citrobacter rodentium* type III secreted protein, EspI, and roles of this and other secreted proteins in infection, *Infect. Immun.*, **72**, 2288–2302.
22. Campellone, K. G., Robbins, D. and Leong, J. M. 2004, EspFU is a translocated EHEC effector that interacts with Tir and N-WASP and promotes Nck-independent actin assembly, *Dev. Cell*, **7**, 217–228.
23. Garmendia, J., Phillips, A. D., Carlier, M. F., et al. 2004, TccP is an enterohaemorrhagic *Escherichia coli* O157:H7 type III effector protein that couples Tir to the actin-cytoskeleton, *Cell Microbiol.*, **6**, 1167–1183.
24. Garmendia, J., Ren, Z., Tennant, S., et al. 2005, Distribution of tccP in clinical enterohemorrhagic and enteropathogenic *Escherichia coli* isolates, *J. Clin. Microbiol.*, **43**, 5715–5720.
25. Garmendia, J. and Frankel, G. 2005, Operon structure and gene expression of the espJ–tccP locus of enterohaemorrhagic *Escherichia coli* O157:H7, *FEMS Microbiol. Lett.*, **247**, 137–145.
26. Sharma, V. K. and Zuerner, R. L. 2004, Role of hha and ler in transcriptional regulation of the esp operon of enterohemorrhagic *Escherichia coli* O157:H7, *J. Bacteriol.*, **186**, 7290–7301.
27. Tatsuno, I., Nagano, K., Taguchi, K., et al. 2003, Increased adherence to Caco-2 cells caused by disruption of the yhiE and yhiF genes in enterohemorrhagic *Escherichia coli* O157:H7, *Infect. Immun.*, **71**, 2598–2606.
28. Iyoda, S. and Watanabe, H. 2005, ClpXP protease controls expression of the type III protein secretion system through regulation of RpoS and GrIR levels in enterohemorrhagic *Escherichia coli*, *J. Bacteriol.*, **187**, 4086–4094.
29. Nadler, C., Shifrin, Y., Nov, S., et al. 2006, Characterization of enteropathogenic *Escherichia coli* mutants that fail to disrupt host cell spreading and attachment to substratum, *Infect. Immun.*, **74**, 839–849.
30. Gomez-Duarte, O. G. and Kaper, J. B. 1995, A plasmid-encoded regulatory region activates chromosomal eaeA expression in enteropathogenic *Escherichia coli*, *Infect. Immun.*, **63**, 1767–1776.
31. Tobe, T., Schoolnik, G. K., Sohel, I., et al. 1996, Cloning and characterization of bfpTVW, genes required for the transcriptional activation of bfpA in enteropathogenic *Escherichia coli*, *Mol. Microbiol.*, **21**, 963–975.
32. Gruenheid, S., Sekirov, I., Thomas, N. A., et al. 2004, Identification and characterization of NleA, a non-LEE-encoded type III translocated virulence factor of enterohaemorrhagic *Escherichia coli* O157:H7, *Mol. Microbiol.*, **51**, 1233–1249.
33. Navarre, W. W., Porwollik, S., Wang, Y., et al. 2006, Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*, *Science*, **313**, 236–238.