

## A 718-kb DNA Sequence of the *Escherichia coli* K-12 Genome Corresponding to the 12.7-28.0 min Region on the Linkage Map

Taku OSHIMA,<sup>1</sup> Hiroji AIBA,<sup>2</sup> Tomoya BABA,<sup>3</sup> Katsutoshi FUJITA,<sup>3</sup> Kouji HAYASHI,<sup>4</sup> Atsuko HONJO,<sup>3</sup> Keiichi IKEMOTO,<sup>3</sup> Toshifumi INADA,<sup>2</sup> Takeshi ITOH,<sup>5</sup> Miwako KAJIHARA,<sup>5</sup> Kiyotaka KANAI,<sup>6</sup> Kaoru KASHIMOTO,<sup>7</sup> Sigenobu KIMURA,<sup>8</sup> Masanari KITAGAWA,<sup>5</sup> Kozo MAKINO,<sup>8</sup> Shinji MASUDA,<sup>6</sup> Takeyoshi MIKI,<sup>7</sup> Kiyoshi MIZOVUCHI,<sup>6</sup> Hirotsada MORI,<sup>5</sup> Kouji MOTOMURA,<sup>6</sup> Yoshikazu NAKAMURA,<sup>1</sup> Hiroko NASHIMOTO,<sup>9</sup> Yoshitaka NISHIO,<sup>6</sup> Noriko SAITO,<sup>9</sup> Gen-ichi SAMPEI,<sup>6</sup> Yasushi SEKI,<sup>6</sup> Hideaki TAGAMI,<sup>2</sup> Keiko TAKEMOTO,<sup>10</sup> Chieko WADA,<sup>10</sup> Yoshihiro YAMAMOTO,<sup>11</sup> Minoru YANO,<sup>5</sup> and Takashi HORIUCHI,<sup>4,\*</sup>

*Institute of Medical Science, University of Tokyo, 4-6-1 Shirogane-dai, Minatoku, Tokyo 108, Japan,<sup>1</sup> Department of Molecular Biology, School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-01, Japan,<sup>2</sup> Department of Biology, Faculty of Science, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe 657, Japan,<sup>3</sup> National Institute for Basic Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki 444, Japan,<sup>4</sup> Research and Education Center of Gene Technology, Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma 630-01, Japan,<sup>5</sup> Department of Applied Physics and Chemistry, University of Electro-Communications, 1-5-1 Chofugaoka, Chofu 182, Japan,<sup>6</sup> Faculty of Pharmaceutical Sciences, Kyushu University, 3-1-1 Maadashi, Higashi-ku, Fukuoka 812-82, Japan,<sup>7</sup> Research Institute for Microbial Diseases, Osaka University, 3-1 Yamadaoka, Suita 562, Japan,<sup>8</sup> Department of BioSciences, Teikyo University, 1-1 Toyosatodai, Utsunomiya 320, Japan,<sup>9</sup> Institute for Virus Research, Kyoto University, Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-01, Japan,<sup>10</sup> and Department of Genetics, Hyogo College of Medicine, 1-1 Mukogawa-cho, Nishinomiya 663, Japan<sup>11</sup>*

(Received 21 June, 1996)

### Abstract

The 718,122 base pair sequence of the *Escherichia coli* K-12 genome corresponding to the region from 12.7 to 28.0 minutes on the genetic map is described. This region contains at least 681 potential open reading frames, of which 277 (41%) have been previously identified, 147 (22%) are homologous to other known genes, 139 (20%) are identical or similar to the hypothetical genes registered in databases, and the remaining 118 (17%) do not show a significant similarity to any other gene. In this region, we assigned a cluster of *cti* genes encoding multienzyme citrate lyase, two clusters of fimbrial genes and a set of lysogenic phage genes encoding integrase, excisionase and repressor in the e14 genetic element. In addition, a new valine tRNA gene, designated *valZ*, and a family of long directly repeated sequences, LDR-A, -B and -C, were found.

**Key words:** *Escherichia coli* K12; genome sequence; ORF analysis

### 1. Introduction

As a model organism, *Escherichia coli* has been playing significant roles in the establishment of a number of basic concepts in molecular biology, and the enormous amount of data accumulated to date has contributed to the understanding of a variety of cellular processes of not only *E. coli* but also other organisms.<sup>1</sup> In addition, determination of the whole genome sequence of *E. coli* is undoubtedly awaited for further insight into various bio-

logical phenomena.

Until now, the complete DNA sequences from two organisms, *Haemophilus influenzae*<sup>2</sup> and *Mycoplasma genitalium*,<sup>3</sup> were reported, and sequencing of several other prokaryotic and eucaryotic genomes are now in progress. In *E. coli*, a contiguous sequence covering about 65% of the whole genome has been determined.<sup>4</sup> The *E. coli* genome project in Japan was started in 1989 to analyze the sequence of the *E. coli* genome from 0 minutes in a clockwise direction. So far, the DNA sequences of the 0-2.1<sup>5</sup> and 2.1-4.1 min<sup>6</sup> regions on the genetic map were determined and published previously. Subsequently, the

Communicated by Mituru Takanami

\* To whom correspondence should be addressed. Tel. +81-564-55-7690, Fax. +81-564-55-7690

entire sequence from 4.1 to 12.7 min was determined and the data of a part of the region (4.1 min to 6.0) registered in the DDBJ database. In April 1995, we reorganized a new group for analysis of the DNA sequence of the region from 13 to 50 minutes using a new protocol. In addition, the team headed by Blattner in the United States has determined a contiguous sequence from 0 (100) to 76 minutes in counterclockwise.<sup>7–12</sup> They have also registered the sequence data of the 76–60 min region in GenBank. In this paper, we describe a 718-kb DNA sequence of the region spanning from 12.7 min to 28.0 min on the genetic map.

## 2. Materials and Methods

### 2.1. DNA source

#### 2.1.1. Preparation of bacterial DNA from Kohara's clones<sup>13</sup>

DNA samples used in this experiment were prepared from Kohara's library, in which DNA fragments of *E. coli* K-12 W3110<sup>14</sup> were cloned in phage lambda. Individual phage clones were grown in 20 ml of NZCYM medium on host W3110, and precipitated by the addition of polyethylene glycol (PEG) #6000. Total phage DNA containing each bacterial DNA fragment was purified with a QIAGEN lambda kit (Qiagen GmbH, Hilden, Germany). Bacterial DNA in each lambda clone was amplified by polymerase chain reaction (PCR) using a GeneAmp XL PCR kit (Perkin-Elmer/Applied Biosystems Division, Norwalk, CT, USA) with two primers, primer 1 (5'-ACAGTCGGTGGTCCGGCAGTAGCAATGGATTACC-3', residing upstream of 678 base pair (bp) from the *Bam*HI cloning site on the left arm of lambda EMBL4) and primer 2 (5'-GCAACCTGCAACCGTATTGAGCGCAAGAATCAGC-3', residing downstream of 640 bp remote from the *Bam*HI cloning site on the right arm).<sup>15</sup> After electrophoresis on a 0.4% agarose gel, the PCR product was purified by EASYTRAP (Takara Shuzo Co., Kyoto, Japan).

#### 2.1.2. Sonication, size selection and cloning of the *E. coli* DNA

Five micrograms of PCR DNA was sonicated for 10–20 sec with a Branson Sonifier (Heat Systems, Farmingdale, NY, USA). The sonicated DNA was blunt-ended using a DNA blunting kit (Takara Shuzo Co.) and fractionated by 6% polyacrylamide gel electrophoresis. DNA fragments of 0.6 ~ 2 kb in length were excised from the gel and eluted by MG elution buffer (500 mM ammonium acetate, 10 mM Magnesium acetate, 1 mM EDTA, 0.1% (W/V) SDS) for overnight at 37°C. The recovered DNA was ligated with the *Sma*I site of vector M13mp19

DNA,<sup>16</sup> which was previously treated with bacterial alkaline phosphatase, and introduced into TG1<sup>17</sup> or JM109<sup>16</sup> competent cells with an electroporator. The cells were plated on LB plates containing X-gal and IPTG to form plaques. More than 250 independent white plaques per lambda clone were isolated.

#### 2.1.3. Purification of the M13 single-stranded DNA

M13 phage clones isolated were grown in 2 × YT medium<sup>17</sup> for 8 hr on host TG1 or JM109. Phage particles were collected from 2 ml of culture supernatant by PEG precipitation and single-stranded DNA (ssDNA) was prepared by phenol-chloroform extraction followed by ethanol precipitation or by an automatic plasmid isolation system (PI-100 Σ; Kurabo, Osaka, Japan).

#### 2.2. DNA sequencing

A DNA sequence kit (Dye primer cycle sequencing ready reaction -21M13; Perkin Elmer/Applied Biosystems Division) was used for Sanger dideoxy chain termination reactions. After termination of the reaction, a portion of the reaction mixtures was applied onto 373S or 377 DNA sequencers (Perkin Elmer/Applied Biosystems Division) and their DNA sequences were analyzed.

#### 2.3. Sequence assembly and data analysis

The raw sequence data were assembled and edited using the two computer programs, Autoassembler of Perkin Elmer/Applied Biosystems Division and ATSQ of Software Co., Ltd., Tokyo, Japan. The programs BLAST and GeneMark<sup>18</sup> were used to aid in identifying potential open reading frames (ORFs). Each ORF was compared to SWISS-PROT and PIR protein databases using BLAST.

#### 2.4. Gap-filling with known sequences

The following four gaps were filled by the sequences registered in GenBank: 6.8-kb region between lambda clones #180 and #203, 1.8 kb between #223 and #225, 1.65 kb between #234 and #236, and 0.38 kb between #173 and #174. The last two clones were reported to be overlapped in the physical map,<sup>4</sup> but we found that there was a gap between them. Other sequences reported in this work were newly determined.

## 3. Results and Discussion

### 3.1. Sequencing strategy

We determined a 718,122-bp sequence spanning the region from 12.7 to 28 min on the *E. coli* genetic map. The strategy for analysis of the *E. coli* genome was based on shotgun cloning and sequencing of the DNA fragments prepared from respective Kohara clones<sup>12</sup> followed by

assembly of random sequences using two computer programs, Autoassembler and ATSQ. Bacterial insert DNA (17 kb on average) in each lambda clone was amplified by a long PCR procedure<sup>15</sup> and PCR products were randomly fragmented by sonication. DNA fragments ranging from 0.6 to 2.0 kb in size were isolated and sub-cloned in phage M13mp19<sup>16</sup> as the vector. ssDNA was purified and used as the template for cycle sequencing with a universal primer (-21M13). When more than 250 random DNA fragments prepared from respective Kohara clones were sequenced, most sequences (98–99%) were assembled into one contiguous sequence. If gaps remained, additional M13 clones or PCR products amplified using appropriately designed synthetic primers encompassing the gaps were sequenced. For ambiguous bases, we sequenced additional M13 clones or PCR products amplified using appropriate M13 clone DNA as a template in reverse direction as well. Random short sequences derived from a Kohara clone were assembled into a contiguous sequence using the Autoassembler program. Then, the contiguous sequences thus obtained from respective Kohara clones were further assembled and edited by the ATSQ program. Under the experimental conditions, mean sequence redundancy was 7.2-fold and the overall accuracy of the contiguous sequence was estimated to be 99.99% by comparison with previously reported sequences and with overlapped sequences of two adjacent lambda clones.

### 3.2. Assignment of potential coding region

To assign ORFs, we first selected potential ORFs which were composed of more than 50 consecutive sense codons. All these ORFs were translated into amino acid sequences and then subjected to homology analysis against the SWISS-PROT (release 31) and PIR (release 44) protein databases using the BLASTP program. When ORFs to be examined showed more than 95% identity in sequence and consistency in amino acid length to known gene products, they were referred to as known genes. In the newly sequenced regions, ORFs with BLASTP scores of more than 100 were classified as either homologous or hypothetical according to their functions. When the functions of the reported ORFs or genes were known, the corresponding new ORFs were classified as homologous; when the functions of the reported ORFs were not known, the corresponding new ORFs were classified as hypothetical. The ORFs that could not be predicted by BLASTP were predicted either using the program GeneMark<sup>18</sup> or visually. In the latter case, the possible NH<sub>2</sub>-terminus of each predicted ORF was estimated by calculating the efficiency of the ribosome binding site for translation with a program developed by ourselves. These ORFs were classified as hypothetical as well. The nomenclature of the predicted ORFs was based on the consecutive numbers of ORFs found in

each Kohara clone; for example, the ORF designated as o162#1 was ORF number 1 in Kohara miniset clone 162. Specific sequences other than ORFs were checked by the FASTA program.

The ORFs predicted in the 12.7–28 min region on the linkage map are summarized in Table 1 and illustrated in the Supplement. We found at least 681 potential ORFs in this region. Among them, 277 ORFs (41%) were identical to known genes reported previously. One hundred and forty seven (22%) and 139 (20%) ORFs were identified as homologous and hypothetical genes, respectively. The rest of the ORFs (118, 17%) did not show a significant similarity to any other gene registered in the two databases.

### 3.3. Unique ORFs identified

In the sequence determined, we identified 9 genes, *nfnB* (nitrofurantoin sensitivity; 13.1 min), *rmk* (regulator of nucleoside diphosphate kinase 14.0), *ghlJ*, *K* and *L* genes (glutamate/aspartate transport system; 14.9 min), *mdaA* (19.4 min), *fskK* (flamentation; 20.3 min), *pgi* (paragut inducible; 22.0 min), *torR* (regulator of trimethylamine oxide reductase; 23.0 min). The DNA or protein sequences of these genes were known but their locations were not identified.

The *lol* (*lpIA*) gene (lipoate-protein ligase) and a part of its sequence had been localized to 99.6 min on the linkage map.<sup>4</sup> However, we found an exactly identical sequence (o214#8) at 20.4 min, an indication that the previous map position of the gene may be erroneous. The deduced amino acid sequence of o207#3 was found to be 70% identical to that of the putative transaldolase encoded by *talC* at 89.1 min. It has been reported that *E. coli* possesses at least two transaldolase genes, *talA* and *talB*, and the amino acid sequence of the *talA* gene product was 64% identical to that of the *talB* gene product.<sup>19</sup> Comparison of amino acid sequence between o207#3 or *TalC* and *TalA* or *TalB*, however, showed only 11 to 14% identity. These results indicate that both o207#3 and *talC* were generated by duplication, and that they belong to a different family from that of *talA* and *talB*.

The amino acid sequences of two ORFs (o162#5 and o162#6) showed a high degree of similarity to *copS* and *copR* gene products of *Pseudomonas syringae*, respectively. Because *CopS* and *CopR* are sensor and transcription activator proteins of the two-component systems, respectively, o162#5 and o162#6 may be new members of the family.

The six ORFs located at 14.0–14.1 min showed a sequence similarity to a group of *cit* genes encoding citrate lyase or its related enzymes in *Klebsiella pneumoniae* and *Haemophilus influenzae*. In *E. coli*, citrate lyase has been shown to be a multienzyme complex built up from at least three different subunits (alpha, beta and

gamma).<sup>20</sup> However, the genes coding for these subunits have not been identified yet. The fact that the order of the six ORFs is identical to that of the *cit* genes in *H. influenzae* strongly suggests that these ORFs are likely to be the *cit* genes in *E. coli*. Moreover, it is interesting to note that a sequence of more than 20 kb covering the region from the *cit*-like gene cluster (14.1 min) to *mrnA* (*pbpA*; 14.5 min) is homologous to a sequence of about 13 kb from the *citG* to *pbp2* genes in *H. influenzae* with respect to genes and gene order, although a central portion of the region present in *E. coli* is not present in *H. influenzae*.

The *cde* gene (constitutive dam expression) is supposed to be a transcription regulator of the *dam* gene encoding DNA adenine methylase. The *cde* gene was reported to be located within the *hlpB* gene (*hpoate* biosynthesis protein; 14.4 min) or between the *hlpB* and *hlpA* genes. Since there is only one ORF (o168#13) present in this spacer, and since it has motifs shared with the LysR transcriptional regulator family, this ORF probably corresponds to the *cde* gene. Likewise, o166#7 was found to have the LysR motifs. This might be the *leuJ* gene encoding a regulator of *leu* and *ile* operons, because it mapped to around this locus.

At least two fimbrial gene clusters were found. One cluster contains three ORFs (o175#9, o175#10 and o175#11; 16.3 min), and the products of the first two genes are homologues of *E. coli* outer membrane chaperone protein PapD and usher protein PapC, respectively. The other contains at least seven fimbrial genes (o221#5–#11; 21.7 min). This cluster also encodes a pair of chaperone- and usher-like proteins.

o225#13 and #14 at 22.8 min probably encode cold shock proteins similar to CspF and CspB proteins, respectively.

A group of ORFs present in a 13-kb region spanning from 24.6 to 24.8 min showed a high degree of similarity in both sequence and organization to the *flag* (Flagella) genes in *Salmonella typhimurium*, except the sequence of the *flagD* gene encoding a flagella basal-body rod modification protein. *E. coli* and *S. typhimurium* seem to have the same set of *flag* genes. In addition, downstream of the *flagMN* genes, there are two other genes (o233#6 and #7) which are almost identical to *Salmonella mivM* and *mivN* genes encoding virulence factors for mouse.

The e14 element (26.1 min) has been thought to be a cryptic prophage. Within the e14 region, three ORFs (o240#2, #3 and #4) are found to be homologues of integrase, excisionase and repressor in lysogenic phage P21. Probably, this putative repressor is inactivated during the process of SOS induction, and consequently, excision of the two genes results in excision of a 14.4-kb circular DNA.

#### 3.4. Other features

We found new long direct repeat sequences, termed LDR-A, -B and -C, at a spacer region between the *kdsA* and *chaA* genes at 27.7 min (Supplement). One repeat unit consists of 535 bp and 2.9 repeat units are present there. At least one homologous unit, termed LDR-D, is present at 73.4 min in *E. coli* and one similar unit is also located at 9 min in *S. typhimurium*.<sup>21</sup>

tRNA genes, IS elements and the RhsC element are located at the sites as reported (Fig. 1). Of the 17 tRNA genes identified, one located between the two *lys*-tRNA genes at 16.9 min seems to encode a new valine tRNA gene, named *valZ*, whose sequence is identical to that of *valT*. Repeated sequences (REP, ERIC and RSA) and *cis*-elements (DNA replication terminus; *Ter* and Chi sequences) are also indicated in Supplement.

**Acknowledgments:** We thank Dr. Y. Kohara for kindly providing all lambda clones used in the present experiment, Dr. N. Ogasawara for giving a valuable advice about the shotgun sequencing protocol and Dr. M. Borodovsky for kindly providing a program software, GeneMark. This work (H. M.) was supported in part by Grants-in-Aid for Scientific Research 04261101 from the Ministry of Education, Science, Sports and Culture, Japan.

#### References

1. Neidhardt, F. C. et al. 1996, *Escherichia coli* and *Salmonella*, American Society for Microbiology, Washington D.C.
2. Fleischmann, R. D., Adams, M. D., White, O. et al. 1995, Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd, *Science*, **269**, 496–512.
3. Fraser, C. M., Gocayne, J. D., White, O. et al. 1995, The minimal gene complement of *Mycoplasma genitalium*, *Science*, **270**, 397–403.
4. Berlyn, M. K. B., Low, B., and Rudd, K. E. 1996, "Linkage map of *Escherichia coli* K-12," 9th ed., in *Escherichia coli* and *Salmonella*, 2nd ed., ed by F. C. Neidhardt, ASM Press, Washington, D.C., pp. 1715–1902.
5. Yura, T., Mori, H., Nagai, H., Nagata, T., Ishihama, A., Fujita, N., Isono, K., Mizobuchi, K., and Nakata, A. 1992, Systematic sequencing of the *Escherichia coli* genome: analysis of the 0–2.4 min region, *Nucleic Acids Res.*, **20**, 3305–3308.
6. Fujita, N., Mori, H., Yura, T., and Ishihama, A. 1994, Systematic sequencing of the *Escherichia coli* genome: analysis of the 2.4–4.1 min (110,917–193,643 bp) region. *Nucleic Acids Res.*, **22**, 1637–1639.
7. Daniels, D. L., Plunkett, G., III, Burland, V., and Blattner, F. R. 1992, Analysis of the *Escherichia coli* genome: DNA sequence of the region from 84.5 to 86.5 minutes, *Science*, **257**, 771–778.
8. Plunkett, G., III, Burland, V., Daniels, D. L., and Blattner, F. R. 1993, Analysis of the *Escherichia coli* genome. III. DNA sequence of the region from 87.2 to

- 89.2 minutes, *Nucleic Acids Res.*, **21**, 3391-3398.
9. Burland, V., Plunkett, G., III, Daniels, D. L., and Blattner, F. R. 1993, DNA sequence and analysis of 136 kilobases of the *Escherichia coli* genome: organizational symmetry around the origin of replication, *Genomics*, **16**, 551-561.
10. Blattner, F. R., Burland, V., Plunkett, G., III, Sofa, H. J., and Daniels, D. L. 1993, Analysis of the *Escherichia coli* genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes, *Nucleic Acids Res.*, **21**, 5408-5417.
11. Sofa, H. J., Burland, V., Daniels, D. L., Plunkett, G., III, and Blattner, F. R. 1994, Analysis of the *Escherichia coli* genome. V. DNA sequence of the region from 76.0 to 81.5 minutes, *Nucleic Acids Res.*, **22**, 2576-2586.
12. Burland, V., Plunkett, G., III, Sofa, H. J., Daniels, D. L., and Blattner, F. R. 1995, Analysis of the *Escherichia coli* genome VI: DNA sequence of the region from 92.8 through 100 minutes, *Nucleic Acids Res.*, **23**, 2105-2119.
13. Kohara, Y., Akiyama, K., and Isono, K. 1987, The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library, *Cell*, **50**, 495-508.
14. Bachmann, B. J. 1972, Pedigrees of some mutant strains of *Escherichia coli* K-12, *Bacteriol. Rev.*, **36**, 525-557.
15. Barnes, W. M. 1994, PCR amplification of up to 35-kb DNA with high fidelity and high yield from  $\lambda$  bacteriophage templates, *Proc. Natl. Acad. Sci. USA*, **91**, 2216-2220.
16. Yanish-Perron, C., Vieira, J., and Messing, J. 1985, Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors, *Gene*, **33**, 103-119.
17. Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989, Molecular cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
18. Borodovsky, M. and McIninch, J. 1993, GENEMARK: Parallel gene recognition for both DNA strands, *Comput. Chem.*, **17**, 123-133.
19. Iida, A., Teshiba, S., and Mizobuchi, K. 1993, Identification and characterization of the *tkfB* gene encoding a second transketolase in *Escherichia coli* K-12, *J. Bacteriol.*, **175**, 5375-5383.
20. Nilekani, S. and SivaRaman, C. 1983, Purification and properties of citrate lyase from *Escherichia coli*, *Biochemistry*, **22**, 4657-4663.
21. Sanderson, K., Hessel, A., Liu, S.-L., and Rudd, K. E. 1996, "The genetic map of *Salmonella typhimurium*," 8th ed., in *Escherichia coli* and *Salmonella*, 2nd ed., ed by F. C. Neidhardt, ASM Press, Washington, D.C., pp. 1903-1999.

Table 1. List of open reading frames of known, suggested or hypothetical genes.

orf number	start (nt)	end (nt)	length (aa)	map (min)	direction <sup>b</sup>	score <sup>c</sup>	% <sup>d</sup>	overlap <sup>e</sup>	gene <sup>f</sup>	y gene <sup>g</sup>	accession <sup>h</sup>	product	species
o162#1	1323	565	253	12.7	-	1604	98.4	254	<i>envY</i>		P10805	SP Porin thermoregulatory protein EnvY.	<i>Escherichia coli</i>
o162#2	2396	1509	296	12.7	-	850	92.9	141	( <i>ybcH</i> )		P37325	SP Hypothetical 16.8 KD protein in <i>nfrA</i> 3' region precursor.	<i>Escherichia coli</i>
o162#3	5369	2400	990	12.7	-	6578	100	990	<i>nfrA</i>		P31600	SP Bacteriophage N4 adsorption protein A.	<i>Escherichia coli</i>
o162#4	7593	5359	745	12.8	-	4942	99.9	745	<i>nfrB</i>		P31599	SP Bacteriophage N4 adsorption protein B precursor.	<i>Escherichia coli</i>
o162#5	9197	7746	484	12.9	-	740	33.9	443	( <i>copS</i> )		Q02541	SP Sensor protein CopS (EC 2.7.3.-).	<i>Pseudomonas syringae</i> (PV. Tomato)
o162#6	9858	9178	227	12.9	-	938	61.3	225	( <i>copR</i> )		Q02540	SP Transcriptional activator protein CopR.	<i>Pseudomonas syringae</i> (PV. Tomato).
o162#7	10015	11385	457	12.9	+	1082	44.5	458	( <i>oprK</i> )		A49937	pir 50k outer membrane protein OprK	<i>Pseudomonas aeruginosa</i>
o162#9	11894	13114	407	13.0	+	182	29.1	206	( <i>mitrC</i> )		P43505	SP Membrane fusion protein MtrC precursor.	<i>Neisseria gonorrhoeae</i>
o163#1	13129	16269	1047	13.0	+	1705	35.5	1058	( <i>czcA</i> )	<i>ybdE</i>	P13511	SP Cation efflux system protein CzcA.	<i>Alcaligenes eutrophus</i>
o163#2	16374	17747	458	13.1	+	3010	99.8	458	<i>pheP</i>		P24207	SP Phenylalanine-specific permease.	<i>Escherichia coli</i>
o163#3	19078	17834	415	13.1	-	486	100	77		<i>ybdG</i>	P39455	SP Hypothetical protein in <i>nfsB</i> 3' region (fragment).	<i>Escherichia coli</i>
o163#4	19839	19189	217	13.1	-	1412	100	217	<i>nfnB, nfsI, nfsB, ntr</i>		P38489	SP Oxygen-insensitive NAD(P)H nitroreductase (EC 1.-.-.-).	<i>Escherichia coli</i>
o163#5	20412	19936	159	13.1	-	221	97.1	35		<i>ybdF</i>	P39454	SP Hypothetical protein in <i>nfsB</i> 5' region (fragment).	<i>Escherichia coli</i>
o163#8	22152	22400	83	13.2	+	134	40.8	49	( <i>srnB</i> )		P13970	SP SmB protein.	<i>Escherichia coli</i>
o163#9	22480	23589	370	13.2	+	2451	100	370	IS186		S40538	pir Hypothetical protein	<i>Escherichia coli</i>
o163#10	23212	22907	102	13.2	-	705	100	102	IS421		S40539	pir Hypothetical protein	<i>Escherichia coli</i>
o163#11	24521	23877	215	13.2	-	1436	99.5	215	<i>entD</i>		JV0078	pir EntD protein	<i>Escherichia coli</i>
o163#12	26963	24672	764	13.2	-	90	100	12	<i>fepA</i>		S06980	pir Ferrienterochelin receptor	<i>Escherichia coli</i>
o164#1	27230	28351	374	13.3	+	2654	100	374	<i>fes</i>		A31958	pir Fes protein	<i>Escherichia coli</i>
o164#2	28357	28572	72	13.3	+	535	100	72		<i>yzfS</i>	A38408	pir Hypothetical protein ( <i>entF</i> 5' region)	<i>Escherichia coli</i>
o164#3	28572	32450	1293	13.3	+	8644	100	1293	<i>entF</i>		B38408	pir Enterochelin synthetase (EC 6.-.-.-) component F	<i>Escherichia coli</i>
o164#4	32670	33800	377	13.4	+	2396	100	377	<i>fepE</i>		P26266	SP Ferric enterobactin transport protein FepE.	<i>Escherichia coli</i>
o164#5	34615	33803	271	13.4	-	1770	100	271	<i>fepC</i>		S16298	pir Ferric enterobactin transport protein FepC	<i>Escherichia coli</i>
o164#6	35604	34615	330	13.4	-	2086	100	330	<i>fepG</i>		S16297	pir Ferric enterobactin transport protein FepG	<i>Escherichia coli</i>
o165#1	36605	35604	334	13.5	-	2061	100	334	<i>fepD</i>		S16296	pir Ferric enterobactin transport protein FepD	<i>Escherichia coli</i>
o165#2	36716	37963	416	13.5	+	2670	100	416		<i>ybdA</i>	S16306	pir Membrane protein P43	<i>Escherichia coli</i>
o165#3	38926	37973	318	13.5	-	1998	99.1	318	<i>fepB</i>		P14609	SP Ferrienterobactin-binding periplasmic protein precursor.	<i>Escherichia coli</i>
o165#4	39136	38933	68	13.5	-	448	98.5	68			JV0044	pir Hypothetical 7.5K protein ( <i>fepB</i> 5' region)	<i>Escherichia coli</i>
o165#5	39301	40473	391	13.6	+	2575	100	391	<i>entC</i>		JT0497	pir Isochorismate synthase (EC 5.4.99.6)	<i>Escherichia coli</i>
o165#6	40486	42093	536	13.6	+	3531	100	536	<i>entE</i>		A48308	pir Enterochelin synthetase (EC 6.-.-.-) component E	<i>Escherichia coli</i>
o165#7	42110	42964	285	13.6	+	1941	100	285	<i>entB, entG</i>		P15048	SP Isochorismatase (EC 3.3.2.1) (2,3-dihydro-2,3-dihydroxybenzoate synthase).	<i>Escherichia coli</i>
o165#8	42967	43710	248	13.6	+	1594	100	248	<i>entA</i>		A91904	pir 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EC 1.3.1.28)	<i>Escherichia coli</i>
o165#9	43716	44126	137	13.6	+	291	42.5	120	( <i>comAB, comA2</i> )	<i>ybdB</i>	P14205	SP ComA operon protein 2.	<i>Bacillus subtilis</i>
o165#10	44310	46412	701	13.7	+	1303	100	200	<i>cstA</i>		P15078	SP Carbon starvation protein A	<i>Escherichia coli</i>
o166#1	46598	46792	65	13.7	+	464	100	65		<i>ybdD</i>	P23518	SP Hypothetical 7.5 KD protein in <i>cstA</i> 3' region.	<i>Escherichia coli</i>
o166#2	47893	46808	362	13.7	-	536	31.8	359	( <i>gldA</i> )	<i>ybdH</i>	P32665	SP Glycerol dehydrogenase (EC 1.1.1.6) (GLDH).	<i>Escherichia coli</i>
o166#3	48002	49021	340	13.7	+	491	26	319			C33496	pir <i>hisC</i> homolog	<i>Bacillus subtilis</i>
o166#7	52033	51134	300	13.8	-	284	23.4	278	( <i>leuO</i> )		P10151	SP Probable activator protein in <i>leuABCD</i> operon.	<i>Escherichia coli</i>
o166#9	53363	53920	186	13.9	+	1235	100	186	<i>ahpC</i>		P26427	SP Alkyl hydroperoxide reductase C22 protein (EC 1.6.4.-) (SCR-23).	<i>Escherichia coli</i>
o166#10	54168	55730	521	13.9	+	1056	100	162	<i>ahpF</i>		P35340	SP Alkyl hydroperoxide reductase F52A protein (EC 1.6.4.-) (fragment).	<i>Escherichia coli</i>
o166#11	56282	55857	142	13.9	-	79	100	12		<i>yzzU</i>	P39177	SP Unknown protein from 2D-PAGE (spots PR25/LM16/2D_000LR3) (fragment).	<i>Escherichia coli</i>
o166#12	56503	57738	412	13.9	+	527	35.5	403	( <i>fdhA</i> )		P46154	SP Glutathione-independent formaldehyde dehydrogenase (EC 1.2.1.46) (FDH) (FALDH).	<i>Pseudomonas putida</i>

Table 1. Continued.

orf number	start* (nt)	end (nt)	length (aa)	map (min)	direction*	score*	%*	overlap*	gene*	y gene*	accession*	product	species
o166#14	58382	57975	136	14.0	-	886	100	136	<i>rnk</i>		P40679	SP Regulator of nucleoside diphosphate kinase.	<i>Escherichia coli</i>
o166#15	59418	58615	268	14.0	-	1832	100	268	<i>rna.rnsA</i>		P21338	SP Ribonuclease I precursor (EC 3.1.27.6) (enterobacter ribonuclease) (RNase I).	<i>Escherichia coli</i>
o166#16	60995	59535	487	14.0	-	528	25.5	479			Q07252	SP Membrane protein.	<i>Alcaligenes eutrophus</i>
o166#17	61924	61049	292	14.0	-	885	51.3	265	( <i>citG</i> )		P45414	SP CitG protein.	<i>Klebsiella pneumoniae</i>
o166#18	62450	61902	183	14.0	-	295	33.1	130	( <i>citG</i> , HI0021)		P44458	SP CitG protein homolog.	<i>Haemophilus influenzae</i>
o167#1	64028	62457	524	14.1	-	2386	72	503	( <i>citF</i> , HI0022)		P44459	SP Citrate lyase alpha chain (EC 4.1.3.6)	<i>Haemophilus influenzae</i>
o167#2	64905	64000	302	14.1	-	1450	76.3	291	( <i>citE</i> , HI0023)		P44460	SP Citrate lyase beta chain (EC 4.1.3.6)	<i>Haemophilus influenzae</i>
o167#3	65198	64905	98	14.1	-	277	48.4	95	( <i>citD</i> )		P02903	SP Citrate lyase gamma chain).	<i>Klebsiella pneumoniae</i>
o167#4	66328	65216	371	14.1	-	1144	52.9	329	( <i>citC</i> , HI0025)		P44462	SP [Citrate (pro-3S)-lyase] ligase (EC 6.2.1.22)	<i>Haemophilus influenzae</i>
o167#5	66650	68305	552	14.1	+	689	27.8	529	( <i>yjdH</i> )		P39272	SP Hypothetical 60.6 KD protein in <i>dcuB-lysU</i> intergenic region (F543).	<i>Escherichia coli</i>
o167#6	68277	68954	226	14.2	+	444	32.6	215			P39486	SP Hypothetical 26.2 KD protein in <i>gdhI</i> 5' region (ORF 1).	<i>Bacillus megaterium</i>
o168#1	69474	69001	158	14.2	-	327	38.8	121	( <i>yhcL</i> )		P45428	SP Hypothetical 48.8 KD protein in <i>nanA-sspB</i> intergenic region (O455).	<i>Escherichia coli</i>
o168#3	69437	70450	338	14.2	+	2267	100	338	IS5		S47725	pir Hypothetical protein F338	<i>Escherichia coli</i>
o168#4	70105	69752	118	14.2	-	778	100	118	IS5		B91483	pir Hypothetical 12K protein (insertion sequence IS5)	<i>Escherichia coli</i>
o168#5	71582	70494	363	14.2	-	705	34.9	327	( <i>yhcL</i> )		P45428	SP Hypothetical 48.8 KD protein in <i>nanA-sspB</i> intergenic region (O455).	<i>Escherichia coli</i>
o168#6	72261	72728	156	14.3	+	1156	100	156		<i>ybeG</i>	P37001	SP Hypothetical 18.4 KD protein in <i>cspE</i> 5' region.	<i>Escherichia coli</i>
o168#7	72906	73112	69	14.3	+	451	100	69	<i>cspE, msmC, gicA</i>		P36997	SP Cold shock-like protein CspE.	<i>Escherichia coli</i>
o168#8	73552	73172	127	14.3	-	834	99.2	127		<i>ybeI</i>	P37002	SP Hypothetical 13.8 KD protein in <i>cspE-lipA</i> intergenic region.	<i>Escherichia coli</i>
o168#9	73645	73869	75	14.3	+	331	100	52		<i>ybeH</i>	P39874	SP Hypothetical protein in <i>cspE-lipA</i> intergenic region (fragment).	<i>Escherichia coli</i>
o168#11	74561	74761	67	14.3	+	408	100	67		<i>ybeC</i>	JN0256	pir hypothetical protein ( <i>lip</i> 3' region)	<i>Escherichia coli</i>
o168#12	75830	74868	321	14.3	-	2168	99.7	321	<i>lipA, lip</i>		P25845	SP Lipoic acid synthetase ( <i>lip-syn</i> ).	<i>Escherichia coli</i>
o168#13	76992	76042	317	14.3	-	351	26.6	271	( <i>leuO</i> )	<i>ybeF</i>	P10151	SP Probable activator protein in <i>leuABCD</i> operon.	<i>Escherichia coli</i>
o168#14	77826	77254	191	14.4	-	1277	99.5	191	<i>lipB</i>		P30976	SP Lipoate biosynthesis protein B.	<i>Escherichia coli</i>
o168#15	78256	77996	87	14.4	-	557	100	87		<i>ybeD</i>	P30977	SP Hypothetical 9.8 KD protein in <i>lipB-dacA</i> intergenic region (ORF1).	<i>Escherichia coli</i>
o168#16	79577	78369	403	14.4	-	2643	100	403	<i>dacA, pfv</i>		P04287	SP Penicillin-binding protein 5 precursor (D-alanyl-D-alanine carboxypeptidase fraction A) (EC 3.4.16.4)	<i>Escherichia coli</i>
o168#17	80804	79719	362	14.4	-	2377	100	362	<i>rlpA</i>		A28387	pir Lipoprotein RlpA precursor	<i>Escherichia coli</i>
o168#18	81927	80818	370	14.5	-	2408	100	370	<i>mrdB, rodA</i>		JT0500	pir Rod shape-determining protein MrdB	<i>Escherichia coli</i>
o169#1	83831	81933	633	14.5	-	4213	100	633	<i>mrDA, pbpA</i>		P08150	SP Penicillin-binding protein 2 (PBP-2).	<i>Escherichia coli</i>
o169#2	84329	83865	155	14.5	-	1043	99.4	155		<i>ybeA</i>	P05850	SP Hypothetical 17.3 KD protein in <i>mrDA/lpbpA</i> 3' region.	<i>Escherichia coli</i>
o169#3	84542	84336	69	14.5	-	451	100	69		<i>ybeB</i>	P05848	SP Hypothetical 7.7 KD protein in <i>mrDA/lpbpA</i> 3' region.	<i>Escherichia coli</i>
o169#4	85524	84913	204	14.5	-	966	70.5	200	( <i>cobC</i> )		P39701	SP Alpha-ribazole-5'-phosphate phosphatase (EC 3.1.3.-).	<i>Salmonella typhimurium</i>
o169#5	86186	85548	213	14.6	-	135	32.3	62			S48587	pir Hypothetical protein	<i>Mycoplasma capricolum</i> (SGC3)
o169#6	87219	86191	343	14.6	-	2245	100	343	<i>holA</i>		A46738	pir DNA-directed DNA polymerase (EC 2.7.7.7) III delta chain	<i>Escherichia coli</i>
o169#7	87800	87222	193	14.6	-	1209	99	193	<i>rlpB</i>		B28387	pir Lipoprotein RlpB precursor	<i>Escherichia coli</i>
o169#8	90397	87818	860	14.6	-	5806	99.7	860	<i>leuS</i>		A30290	pir Leucine-tRNA ligase (EC 6.1.1.4)	<i>Escherichia coli</i>
o169#9	90632	91111	160	14.7	+	516	100	85		<i>ybeL</i>	P46129	SP Hypothetical protein in <i>leuS</i> 3' region (fragment).	<i>Escherichia coli</i>
o169#10	92167	91187	327	14.7	-	183	35.2	88			P34226	SP Skt5 protein.	<i>Saccharomyces cerevisiae</i> (Baker's Yeast)
o169#13	95020	94469	184	14.7	-	170	32.4	102			S50869	pir Hypothetical protein 1	<i>Pseudomonas aeruginosa</i>
o170#1	99007	97340	556	14.8	-	703	33.5	568			P07823	SP 78 KD glucose regulated protein precursor (GRP 78) (BIP).	<i>Mesocricetus auratus</i> (Golden hamster)

Table 1. Continued.

orf number	start* (nt)	end (nt)	length (aa)	map (min)	direction <sup>b</sup>	score <sup>c</sup>	% <sup>d</sup>	overlap <sup>e</sup>	gene <sup>f</sup>	y gene <sup>g</sup>	accession <sup>h</sup>	product	species	
o170#2	100026	99094	311	14.8	-	788	100	123		<i>ybeK</i>	P41409	SP	Hypothetical protein in <i>gltL</i> 3' region (fragment).	<i>Escherichia coli</i>
o170#3	100869	100147	241	14.9	-	1561	100	241	<i>gltL</i>		P41076	SP	Glutamate/aspartate transport ATP-binding protein GltL.	<i>Escherichia coli</i>
o170#4	101543	100872	224	14.9	-	1413	100	224	<i>gltK</i>		P41075	SP	Glutamate/aspartate transport system permease protein GltK.	<i>Escherichia coli</i>
o170#5	102283	101546	246	14.9	-	1633	100	246	<i>gltJ</i>		P41074	SP	Glutamate/aspartate transport system permease protein GltJ.	<i>Escherichia coli</i>
o170#6	103436	102456	327	14.9	-	327	30.8	234	<i>(glnH)</i>	<i>ybeJ</i>	P27676	SP	Glutamine-binding protein precursor.	<i>Bacillus stearothermophilus</i>
o170#7	104627	103614	338	14.9	-	2267	100	338	IS5		S47725	pir	Hypothetical protein F338	<i>Escherichia coli</i>
o170#8	103959	104312	118	15.0	+	778	100	118	IS5		B91483	pir	Hypothetical 12KD protein (insertion sequence IS5)	<i>Escherichia coli</i>
o170#9	106495	104960	512	15.0	-	3431	99.8	512	<i>Int, cutE</i>		P23930	SP	Apolipoprotein N-acyltransferase (EC 2.3.1.-) (alp N-acyltransferase) (copper homeostasis protein CutE).	<i>Escherichia coli</i>
o170#10	107215	106523	231	15.0	-	338	23.9	264	<i>(yffL)</i>		P39319	SP	Hypothetical 49.8 KD protein in <i>cysQ-msrA</i> intergenic region (F447).	<i>Escherichia coli</i>
o171#1	107955	107491	155	15.0	-	156	35.9	92			C36933	pir	Orf 5' of diacylglycerol kinase homolog	<i>Streptococcus mutans</i>
o171#2	109031	107955	359	15.0	-	592	41.4	232	<i>(phoH, psiH)</i>		P31544	SP	PhoH protein (phosphate starvation-inducible protein PsiH).	<i>Escherichia coli</i>
o171#3	109819	109148	224	15.1	-	818	63.9	194			S49379	pir	Hypothetical protein 2	<i>Pseudomonas aeruginosa</i>
o171#4	110568	109795	258	15.1	-	1224	70.3	246			S49379	pir	Hypothetical protein 2	<i>Pseudomonas aeruginosa</i>
o171#5	110714	111886	391	15.1	+	853	38.5	392	<i>(visC)</i>		P25535	SP	VisC protein.	<i>Escherichia coli</i>
o171#6	114789	113128	554	15.2	-	3735	100	554	<i>asnB</i>		A36616	pir	Asparagine synthase (glutamine-hydrolyzing) (EC 6.3.5.4)	<i>Escherichia coli</i>
o171#7	115938	115189	250	15.2	-	1668	100	250	<i>nagD</i>		S06991	pir	NagD protein	<i>Escherichia coli</i>
o171#8	117206	115989	406	15.2	-	2623	100	406	<i>nagC</i>		C37018	pir	NagC protein	<i>Escherichia coli</i>
o171#9	118363	117218	382	15.2	-	2481	100	382	<i>nagA</i>		S06989	pir	N-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25)	<i>Escherichia coli</i>
o171#10	119223	118426	266	15.3	-	1772	100	266	<i>nagB, glmD</i>		P09375	SP	Glucosamine-6-phosphate isomerase (EC 5.3.1.10) (glucosamine-6-phosphate deaminase).	<i>Escherichia coli</i>
o172#1	119556	121499	648	15.3	+	4169	100	648	<i>nagE, pstN</i>		P09323	SP	(Phosphotransferase enzyme II,ABC component) (EC 2.7.1.69)	<i>Escherichia coli</i>
o172#2	121708	123366	553	15.3	+	3757	100	553	<i>glnS</i>		P00962	SP	Glutamyl-tRNA synthetase (EC 6.1.1.18)	<i>Escherichia coli</i>
o172#5	126257	125814	148	15.4	-	1000	100	148	<i>fur</i>		S07308	pir	Regulatory protein Fur	<i>Escherichia coli</i>
o172#6	126302	126553	84	15.4	+	556	98.8	84		<i>ybfJ</i>	P46146	SP	Hypothetical 10.2 KD protein in <i>fur-flaA</i> intergenic region	<i>Escherichia coli</i>
o172#7	127076	126549	176	15.4	-	1202	100	176	<i>fldA</i>		A37319	pir	Flavodoxin	<i>Escherichia coli</i>
o172#10	128598	129140	181	15.5	+	1154	100	181	<i>seqA</i>		A54296	pir	SeqA protein	<i>Escherichia coli</i>
o172#11	129169	130806	546	15.5	+	3614	100	546	<i>pgm</i>		P36938	SP	Phosphoglucosmutase (EC 5.4.2.2)	<i>Escherichia coli</i>
o172#13	131920	131561	120	15.6	-	771	99.1	113		<i>ybfG</i>	P37003	SP	Hypothetical 13.2 KD protein in <i>pgm-potE</i> intergenic region.	<i>Escherichia coli</i>
o173#1	132316	132002	105	15.6	-	683	100	105		<i>ybfH</i>	P37004	SP	Hypothetical 11.9 KD protein in <i>pgm-potE</i> intergenic region.	<i>Escherichia coli</i>
o173#2	133876	132560	439	15.6	-	2841	100	439	<i>potE</i>		P24170	SP	Putrescine-ornithine antiporter (putrescine transport protein).	<i>Escherichia coli</i>
o173#3	136071	133876	732	15.6	-	4991	100	732	<i>speF</i>		P24169	SP	Ornithine decarboxylase, inducible (EC 4.1.1.17).	<i>Escherichia coli</i>
o173#4	136194	136448	85	15.7	+	574	100	85		<i>ybfK</i>	P46121	SP	Hypothetical 9.5 KD protein in <i>speF-kdpE</i> intergenic region.	<i>Escherichia coli</i>
o173#5	137269	136670	200	15.7	-	1336	92.4	225	<i>kdpE</i>		C42372	pir	Regulatory protein KdpE	<i>Escherichia coli</i>
o173#6	140025	137344	894	15.7	-	5863	100	894	<i>kdpD</i>		B42372	pir	Regulatory protein KdpD	<i>Escherichia coli</i>
o173#7	140590	140021	190	15.7	-	1231	100	190	<i>kdpC</i>		P03961	SP	Potassium-transporting ATPase (EC 3.6.1.36), C chain (ATP phosphohydrolase (potassium-transporting), C chain).	<i>Escherichia coli</i>
o173-174#1	142647	140602	682	15.8	-	4297	99.9	682	<i>kdpB</i>		P03960	SP	Potassium-transporting ATPase (EC 3.6.1.36), B chain (ATP phosphohydrolase (potassium-transporting), B chain).	<i>Escherichia coli</i>
o174#1	144343	142673	557	15.8	-	3608	100	557	<i>kdpA</i>		P03959	SP	Potassium-transporting ATPase (EC 3.6.1.36), A chain (ATP phosphohydrolase (potassium-transporting), A chain).	<i>Escherichia coli</i>
o174#2	144432	144346	29	15.8	-	165	96.6	29	<i>kdpF</i>		P36937	SP	KdpF protein.	<i>Escherichia coli</i>
o174#3	144745	144948	68	15.8	+	478	100	68		<i>ybfA</i>	P28913	SP	Hypothetical 8.3 KD protein in <i>rhcS</i> 5' region precursor.	<i>Escherichia coli</i>
o174#4	145194	149384	1397	15.9	+	9653	99.9	1397	<i>rhcC</i>		P16918	SP	RhcC protein precursor.	<i>Escherichia coli</i>
o174#5	149387	149710	108	15.9	+	694	100	108		<i>ybfB</i>	P28914	SP	Hypothetical 12.6 KD protein in <i>rhcC</i> 3' region (ORF-C2).	<i>Escherichia coli</i>
o174#6	149831	151261	477	16.0	+	2397	96.6	354	<i>rhcA</i>		P16916	SP	RhcA protein precursor.	<i>Escherichia coli</i>



Table 1. Continued.

orf number	start (nt)	end (nt)	length (aa)	map (min)	direction <sup>b</sup>	score <sup>c</sup>	% <sup>d</sup>	overlap <sup>e</sup>	gene <sup>f</sup>	y gene <sup>g</sup>	accession <sup>h</sup>	product	species
o174#7	151261	151827	189	16.0	+	1258	100	189		<i>ybfC</i>	P28915	SP Hypothetical 22.0 KD protein in <i>rhcC</i> 3' region precursor (ORF-C3).	<i>Escherichia coli</i>
o174#8	152056	152307	84	16.0	+	507	96.2	79		<i>ybfD</i>	P28916	SP H repeat-associated protein in <i>rhcC</i> 3' region (ORF-H3).	<i>Escherichia coli</i>
o174#9	152511	153569	353	16.0	+	184	63.6	44 ( <i>rfbQ</i> )		<i>yhhI</i>	S28481	PIR RfbQ protein	<i>Vibrio cholerae</i>
o174#10	153703	154461	253	16.0	+	595	52	173 ( <i>rfbR</i> )		<i>ybfD</i>	S28482	PIR RfbR protein	<i>Vibrio cholerae</i>
o174#11	154612	155118	169	16.1	+	33	33.3	69 ( <i>rfbR</i> )		<i>ybgA</i>	S28482	PIR RfbR protein	<i>Vibrio cholerae</i>
o174#12	155118	156533	472	16.1	+	3246	100	472 <i>phr</i>			P00914	SP Deoxyribodipyrimidine photolyase (EC 4.1.99.3)	<i>Escherichia coli</i>
o175#1	158167	156689	493	16.1	-	463	30.4	484 ( <i>dtpT</i> )			P36574	SP Di-tripeptide transporter.	<i>Lactococcus lactis</i> (subsp. <i>lactis</i> ) ( <i>Streptococcus lactis</i> )
o175#2	158438	159178	247	16.1	+	409	50	126			P45371	SP Hypothetical protein in <i>phbB</i> 3' region (ORF7) (fragment).	<i>Chromatium vinosum</i>
o175#3	159204	159857	218	16.2	+	685	47.4	209 (HI1731)			P44299	SP Hypothetical protein HI1731.	<i>Haemophilus influenzae</i>
o175#4	159854	160783	310	16.2	+	323	30.9	320 (DUR1.2, YBR208C, YBR1448)			P32528	SP Urea amidolyase (contains: urea carboxylase (EC 6.3.4.6) /Allophanate hydrolase (EC 3.5.1.54)).	<i>Saccharomyces cerevisiae</i> (baker's yeast).
o175#5	160776	161507	244	16.2	+	376	44.9	127 ( <i>lamB</i> )			P38096	SP Lactam utilization protein LamB.	<i>Emericella nidulans</i> ( <i>Aspergillus nidulans</i> )
o175#9	165318	164593	242	16.3	-	604	39.7	242 ( <i>papD</i> )			P15319	SP Chaperone protein PapD precursor.	<i>Escherichia coli</i>
o175#10	167831	165336	832	16.3	-	1642	36.9	835 ( <i>papC</i> )			P07110	SP Outer membrane usher protein PapC precursor.	<i>Escherichia coli</i>
o175#11	168406	167843	188	16.3	-	237	27.8	187 ( <i>papA</i> )			A23221	PIR Pap fimbrial protein precursor	<i>Escherichia coli</i>
o176#1	170079	168799	427	16.4	-	2837	99.3	427 <i>glfA, glfT, icdB</i>		<i>ybgD</i>	P00891	SP Citrate synthase (EC 4.1.3.7).	<i>Escherichia coli</i>
o176#2	170788	171174	129	16.4	+	818	100	129 <i>sdhC, cybA</i>			P10446	SP Succinate dehydrogenase cytochrome B-556 subunit.	<i>Escherichia coli</i>
o176#3	171171	171515	115	16.4	+	731	100	115 <i>sdhD</i>			B28836	PIR Succinate dehydrogenase (EC 1.3.99.1) 13K hydrophobic protein	<i>Escherichia coli</i>
o176#4	171518	173281	588	16.4	+	3935	99.5	588 <i>sdhA</i>			C28836	PIR Succinate dehydrogenase (EC 1.3.99.1) flavoprotein	<i>Escherichia coli</i>
o176#5	173300	174013	238	16.5	+	1664	100	238 <i>sdhB</i>			A28837	PIR Succinate dehydrogenase (EC 1.3.99.1) iron-sulfur protein	<i>Escherichia coli</i>
o176#6	174317	177115	933	16.5	+	6280	99.9	933 <i>sucA</i>			A30256	PIR Oxoglutarate dehydrogenase (lipoamide) (EC 1.2.4.2)	<i>Escherichia coli</i>
o176#7	177133	178347	405	16.5	+	2561	100	405 <i>sucB</i>			P07016	SP Dihydroliipoamide succinyltransferase component (E2) of 2-oxoglutarate dehydrogenase complex (EC 2.3.1.61).	<i>Escherichia coli</i>
o176#8	178625	179788	388	16.6	+	2519	100	388 <i>sucC</i>			A24090	PIR succinate-CoA ligase (ADP-forming) (EC 6.2.1.5) beta chain	<i>Escherichia coli</i>
o176#9	179791	180657	289	16.6	+	1874	100	289 <i>sucD</i>			A90499	PIR succinate-CoA ligase (ADP-forming) (EC 6.2.1.5) alpha chain	<i>Escherichia coli</i>
o177#1	181486	180767	240	16.6	-	1568	100	240 <i>farR, g30</i>			P13669	SP Fatty acyl responsive regulator (P30 protein).	<i>Escherichia coli</i>
o177#2	181595	183568	658	16.6	+	1009	34.2	579 ( <i>fruA</i> )			P23387	SP PTS system.(phosphotransferase enzyme II, BC component) (EC 2.7.1.69) (EII-FRU).	<i>Rhodobacter capsulatus</i> ( <i>Rhodospseudomonas capsulata</i> )
o178#2	187069	188634	522	16.8	+	3417	99.8	522 <i>cydA, cyd-1</i>			P11026	SP Cytochrome d ubiquinol oxidase subunit I (EC 1.10.3.-).	<i>Escherichia coli</i>
o178#3	188653	189789	379	16.8	+	2543	100	379 <i>cydB, cyd-2</i>			P11027	SP Cytochrome d ubiquinol oxidase subunit II (EC 1.10.3.-).	<i>Escherichia coli</i>
o178#5	190363	190764	134	16.8	+	860	99.3	134		<i>ybgC</i>	P08999	SP Hypothetical 15.6 KD protein in <i>cydB-tolQ</i> intergenic region.	<i>Escherichia coli</i>
o178#6	190764	191453	230	16.8	+	1450	99.6	230 <i>tolQ, fii</i>			B25980	PIR TolQ protein	<i>Escherichia coli</i>
o178#7	191460	191885	142	16.9	+	886	100	142 <i>tolR</i>			C25980	PIR TolR protein	<i>Escherichia coli</i>
o178#8	191953	193215	421	16.9	+	2537	99.8	421 <i>tolA, cim, excC, lky</i>			P19934	SP TolA protein.	<i>Escherichia coli</i>
o178#9	193348	194640	431	16.9	+	2877	100	431 <i>tolB</i>			JV0058	PIR TolB protein	<i>Escherichia coli</i>
o178#10	194678	195196	173	16.9	+	1120	100	173 <i>pal, excC</i>			P07176	SP Peptidoglycan-associated lipoprotein precursor.	<i>Escherichia coli</i>
o178#11	195209	195997	263	16.9	+	1368	92	237 ( <i>ybgF</i> )			P45955	SP Hypothetical 28.2 KD protein in <i>pal-lysT</i> intergenic region.	<i>Escherichia coli</i>
o178#12	197696	198736	347	17.0	+	2152	97.9	330 <i>nadA, nicA</i>			P11458	SP Quinolinate synthetase A.	<i>Escherichia coli</i>
o179#1	198777	199493	239	17.0	+	420	95.5	66 <i>pnuC</i>			P31215	SP PnuC protein (fragment).	<i>Escherichia coli</i>
o179#2	200434	199496	313	17.0	-	513	43.6	202 ( <i>cxcD</i> )			P13512	SP Cation efflux system protein CzcD.	<i>Alcaligenes eutrophus</i>
o179#5	201244	202293	350	17.1	+	2316	100	350 <i>aroG</i>			P00886	SP Phospho-2-dehydro-3-deoxyheptonate aldolase.phe-sensitive (EC 4.1.2.15)	<i>Escherichia coli</i>
o179#6	203203	202457	249	17.1	-	1647	100	249 <i>xpmA, xpm</i>			P31217	SP Phosphoglycerate mutase I (EC 5.4.2.1)	<i>Escherichia coli</i>

Table 1. Continued.

orf number	start* (nt)	end (nt)	length (aa)	map (min)	direction*	score*	%*	overlap* gene*	y gene*	accession*	product	species
o179#7	204448	203411	346	17.1	-	2351	99.7	346 <i>galM</i>		P40681	SP Aldose 1-epimerase (EC 5.1.3.3)	<i>Escherichia coli</i>
o179#8	205590	204445	382	17.1	-	2511	100	382 <i>galK</i>		B23044	pir Galactokinase (EC 2.7.1.6)	<i>Escherichia coli</i>
o179#9	206640	205597	348	17.2	-	2432	100	348 <i>galT</i>		S00722	pir UDPglucose-hexose-1-phosphate uridylyltransferase (EC	<i>Escherichia coli</i>
o179#10	207666	206653	338	17.2	-	2307	100	338 <i>galE, galD</i>		P09147	SP UDP-glucose 4-epimerase (EC 5.1.3.2)	<i>Escherichia coli</i>
o179#11	208949	207930	340	17.2	-	2222	99.1	340 <i>phrA</i>		P31060	SP Photorepair protein PhrA.	<i>Escherichia coli</i>
o180#1	210255	209470	262	17.2	-	450	37.6	245 ( <i>modA</i> )		P37733	SP Molybdenum transport protein ModA.	<i>Azotobacter vinelandii</i>
o180#3	210700	211470	257	17.3	+	1634	100	257 <i>modA</i>		P37329	SP Molybdate-binding periplasmic protein precursor.	<i>Escherichia coli</i>
o180#4	211473	212159	229	17.3	+	1475	99.6	229 <i>modB, chlJ</i>		P09834	SP Molybdenum transport system permease protein ModB.	<i>Escherichia coli</i>
o180#5	212165	213220	352	17.3	+	2293	100	352 <i>modC, chlD, narD</i>		P09833	SP Molybdenum transport ATP-binding protein ModC.	<i>Escherichia coli</i>
o180#6	213075	213212	46	17.3	+	243	100	32 <i>modD</i>		P21829	SP Putative molybdenum transport protein ModD (fragment).	<i>Escherichia coli</i>
o180#7	214042	213227	272	17.3	-	173	29.6	287 ( <i>cof</i> )		P46891	SP Cof protein.	<i>Escherichia coli</i>
o180#8	214197	215189	331	17.4	+	223	27.2	169		P46057	SP Muconate cycloisomerase I (EC 5.5.1.1)	<i>Trichosporon cutaneum</i>
o180#9	216249	215236	338	17.4	-	400	28.2	287 ( <i>oxyR, momR, mor</i> )		P11721	SP Hydrogen peroxide-inducible genes activator (morphology and auto-aggregation control protein).	<i>Escherichia coli</i>
o180#11	217498	218928	477	17.4	+	512	26.5	486		Q07252	SP Membrane protein	<i>Alcaligenes eutrophus</i>
o180#13	219479	221371	631	17.5	+	348	25.1	690		P20004	SP Aconitate hydratase (EC 4.2.1.3)	<i>Bos taurus</i> (Bovine)
o201#1	222891	221611	427	17.5	-	1955	98	300	<i>ybhC</i>	P46130	SP Hypothetical protein in <i>bioA</i> 5' region (fragment).	<i>Escherichia coli</i>
o201#2	223519	223046	158	17.5	-	1102	99.4	158	<i>ybhB</i>	P12994	SP Hypothetical 17.1 KD protein in <i>bioA</i> 5' region.	<i>Escherichia coli</i>
o201#3	224870	223581	430	17.6	-	2929	100	430 <i>bioA</i>		A32025	pir Adenosylmethionine-8-amino-7-oxononanoate transaminase (EC 2.6.1.62)	<i>Escherichia coli</i>
o201#4	224957	225994	346	17.6	+	2297	100	346 <i>bioB</i>		C32025	pir Biotin synthetase	<i>Escherichia coli</i>
o201#5	225994	227145	384	17.6	+	2536	100	384 <i>bioF</i>		P12998	SP 8-amino-7-oxononanoate synthase (EC 2.3.1.47)	<i>Escherichia coli</i>
o201#6	227135	227887	251	17.6	+	1718	100	251 <i>bioC</i>		E32025	pir BioC protein	<i>Escherichia coli</i>
o201#7	227886	228557	224	17.6	+	1468	100	224 <i>bioD</i>		P13000	SP Dethiobiotin synthetase (EC 6.3.3.3)	<i>Escherichia coli</i>
o201#8	229138	231156	673	17.7	+	4330	100	673 <i>uvrB</i>		A93613	pir UvrB protein	<i>Escherichia coli</i>
o203#1	232259	231354	302	17.7	-	448	35.4	229		S27544	pir Hypothetical protein 1	<i>Thermoanaerobacterium thermosulfurigenes</i>
o203#2	232656	233642	329	17.8	+	2204	100	329 <i>moaA, chlA1, chlA, narA, bisA</i>		P30745	SP Molybdenum cofactor biosynthesis protein A.	<i>Escherichia coli</i>
o203#3	233667	234176	170	17.8	+	1109	100	170 <i>moaB</i>		S34999	pir MoaB protein	<i>Escherichia coli</i>
o203#4	234182	234664	161	17.8	+	962	95.7	161 <i>moaC</i>		S31881	pir MoaC protein	<i>Escherichia coli</i>
o203#5	234660	234902	81	17.8	+	519	98.8	81 <i>moaD, chlA4, chlM</i>		P30748	SP Molybdopterin (mpt) converting factor, subunit 1 (molybdenum cofactor biosynthesis protein D).	<i>Escherichia coli</i>
o203#6	234907	235356	150	17.8	+	1025	100	150 <i>moaE</i>		S35002	pir MoaE protein	<i>Escherichia coli</i>
o203#7	235496	236197	234	17.8	+	168	22.5	204 ( <i>yccA</i> )		P06967	SP Hypothetical 23.4 KD protein in <i>serT</i> 5' region.	<i>Escherichia coli</i>
o203#10	239351	238113	413	17.9	-	409	27.2	372 ( <i>cls, nov</i> )		P31071	SP Cardiolipin synthetase (EC 2.7.8.-).	<i>Escherichia coli</i>
o203#13	241177	240617	187	17.9	-	401	25.2	290 ( <i>yhhJ</i> )		P31993	SP Hypothetical 41.1 KD protein in <i>rhsB-pit</i> intergenic region.	<i>Escherichia coli</i>
o203#15	242863	241733	377	18.0	-	609	31.5	375 ( <i>yhiG</i> )		P37624	SP Hypothetical ABC transporter in <i>rhsB-pit</i> intergenic region (F648).	<i>Escherichia coli</i>
o203#16	244607	242859	583	18.0	-	412	34.7	239 ( <i>abc2</i> )		P41234	SP ATP-binding cassette transporter 2 (fragment).	<i>Mus musculus</i> (Mouse)
o203#17	245583	244588	332	18.0	-	297	25.5	274 ( <i>emrA, HI0898</i> )		P44928	SP Multidrug resistance protein a homolog.	<i>Haemophilus influenzae</i>
o203#18	246254	245586	223	18.0	-	677	100	103	<i>ybhH</i>	P41037	SP Hypothetical transcriptional regulator in <i>moaE-rhlE</i> intergenic region (fragment).	<i>Escherichia coli</i>
o203#19	246483	247844	454	18.1	+	2970	100	454 <i>rhlE</i>		P25888	SP Putative ATP-dependent RNA helicase RhlE.	<i>Escherichia coli</i>
o203#20	248561	248082	160	18.1	-	1077	100	160	<i>ybiA</i>	P30176	SP Hypothetical 18.7 KD protein in <i>rhlE-dinG/rarB</i> intergenic region (F160).	<i>Escherichia coli</i>
o204#1	248681	250828	716	18.1	+	4732	100	716 <i>dinG, rarB</i>		P27296	SP Probable ATP-dependent helicase DinG (DNA-damage-inducible protein G).	<i>Escherichia coli</i>
o204#2	250859	251818	320	18.1	+	2137	100	320	<i>ybiB</i>	P30177	SP Hypothetical 35.0 KD protein in <i>dinG/rarB</i> 3' region (O320).	<i>Escherichia coli</i>
o204#3	251962	253044	361	18.2	+	676	37.5	323 ( <i>ldh</i> )	<i>ybiC</i>	Q07251	SP L-lactate dehydrogenase (EC 1.1.1.27).	<i>Alcaligenes eutrophus</i>

Table 1. Continued.

orf number	start* (nt)	end (nt)	length (aa)	map (min)	direction*	score*	%*	overlap* gene'	y gene'	accession <sup>b</sup>	product	species	
o204#4	253536	253279	86	18.2	-	502	100	86	<i>ybiJ</i>	P41038	SP	Hypothetical 8.6 KD protein in <i>dinG/rarB</i> 3' region.	<i>Escherichia coli</i>
o204#5	254067	253804	88	18.2	-	175	42.4	66 ( <i>traR</i> )	<i>ybiI</i>	P41065	SP	TraR protein.	<i>Escherichia coli</i>
o204#7	257142	254863	760	18.2	-	351	23.1	789 ( <i>fptA</i> )		P42512	SP	Fe(III)-pyochelin receptor precursor.	<i>Pseudomonas</i>
o204#8	257811	257410	134	18.3	-	223	46.5	86 ( <i>ybiJ</i> )		P41038	SP	Hypothetical 8.6 KD protein in <i>dinG/rarB</i> 3' region.	<i>Escherichia coli</i>
o205#1	260809	258869	647	18.3	-	187	22.3	337 ( <i>yjeP</i> )		P39285	SP	Hypothetical 123.8 KD protein in <i>genX-psd</i> intergenic region precursor (F1107).	<i>Escherichia coli</i>
o205#3	262073	261354	240	18.4	-	1542	99.6	240 <i>glnQ</i>		S03183	PIR	Glutamine transport protein GlnQ	<i>Escherichia coli</i>
o205#4	262729	262073	219	18.4	-	1409	100	219 <i>glnP</i>		S03182	PIR	Glutamine transport protein GlnP	<i>Escherichia coli</i>
o205#5	263614	262871	248	18.4	-	1586	100	248 <i>glnH</i>		S03181	PIR	Glutamine-binding protein precursor	<i>Escherichia coli</i>
o205#6	264521	264021	167	18.4	-	1054	100	167 <i>dps</i>		S29942	PIR	Dps protein	<i>Escherichia coli</i>
o205#7	265707	264823	295	18.5	-	1883	99.7	295	<i>ybiF</i>	P36545	SP	Hypothetical 31.3 KD protein in <i>dps/pexB</i> 5' region (ORF1).	<i>Escherichia coli</i>
o205#8	266060	266572	171	18.5	+	1131	100	171 <i>ompX</i>		P36546	SP	Outer membrane protein X precursor.	<i>Escherichia coli</i>
o205#9	268207	266627	527	18.5	-	882	32.3	495 (HI1005)		P44974	SP	Hypothetical protein HI1005.	<i>Haemophilus influenzae</i>
o206#1	269257	270372	372	18.5	+	77	23.3	408 ( <i>ag45</i> )		P46838	SP	46 KD membrane protein.	<i>Mycobacterium leprae</i>
o206#2	271354	270437	306	18.6	-	1300	64.4	295 ( <i>yeeG</i> , <i>erfK</i> )		P39176	SP	31.6 KD protein in <i>cobT</i> 3' region precursor.	<i>Escherichia coli</i>
o206#3	271573	273162	530	18.6	+	260	27.8	439 ( <i>tef3</i> )		P29551	SP	Elongation factor 3 (EF-3).	<i>Pneumocystis carinii</i>
o206#5	275638	274826	271	18.7	-	157	36.6	71		P21878	SP	Hypothetical protein in <i>pdhA</i> 5' region (ORF1) (fragment).	<i>Bacillus stearothermophilus</i>
o207#1	278216	275787	810	18.7	-	767	38	806 ( <i>pfID</i> )		P32674	SP	Formate acetyltransferase 2 (EC 2.3.1.54)	<i>Escherichia coli</i>
o207#2	279148	278225	308	18.7	-	390	35.4	308 ( <i>pfIC</i> )		P32675	SP	Probable pyruvate formate-lyase 2 activating enzyme (EC 1.97.1.4).	<i>Escherichia coli</i>
o207#3	279180	279911	244	18.8	+	998	70	220 ( <i>talC</i> )		P32669	SP	Transaldolase-like protein (EC 2.2.1.-).	<i>Escherichia coli</i>
o207#4	280739	279993	249	18.8	-	1634	100	249 <i>moeB</i> , <i>chlN</i>		P12282	SP	Molybdopterin biosynthesis MoeB protein.	<i>Escherichia coli</i>
o207#5	281974	280742	411	18.8	-	2711	100	411 <i>moeA</i> , <i>chlE</i> , <i>bisB</i> , <i>narE</i>		P12281	SP	Molybdopterin biosynthesis MoeA protein.	<i>Escherichia coli</i>
o207#6	282178	283140	321	18.8	+	314	100	52	<i>ybiK</i>	P37595	SP	Hypothetical protein in <i>moeA-grxA</i> intergenic region precursor (fragment).	<i>Escherichia coli</i>
o207#7	283022	284998	659	18.8	+	798	48.1	557 ( <i>yejF</i> )		P33916	SP	Hypothetical ABC transporter in <i>bcr</i> 5' region.	<i>Escherichia coli</i>
o208#1	284763	286556	598	18.9	+	624	30.6	481 ( <i>hbpA</i> , <i>dppA</i> , HI0853)		P33950	SP	Heme-binding protein A precursor (Hemin-binding LipO protein).	<i>Haemophilus influenzae</i>
o208#2	286577	287494	306	18.9	+	804	41.1	309 ( <i>nikB</i> )		S47696	PIR	Hypothetical protein O314	<i>Escherichia coli</i>
o208#3	287500	288408	303	18.9	+	814	42	283 ( <i>dppC</i> )		P37315	SP	Dipeptide transport system permease protein DppC.	<i>Escherichia coli</i>
o208#4	288589	290934	782	19.0	+	449	31.1	309		Q04855	SP	Hypothetical 80.5 KD protein in <i>ntrC</i> 5' region (ORF1).	<i>Azorhizobium caulinodans</i>
o208#6	293645	292323	441	19.1	-	560	27.4	445		S49379	PIR	Hypothetical protein 2	<i>Pseudomonas aeruginosa</i>
o209#3	296337	297536	400	19.1	+	2604	99.8	400 <i>dacC</i>		P08506	SP	Penicillin-binding protein 6 precursor (D-alanyl-D-alanine carboxypeptidase fraction C) (EC 3.4.16.4)	<i>Escherichia coli</i>
o209#4	298344	297589	252	19.2	-	1689	100	252 <i>deoR</i> , <i>nucR</i>		P06217	SP	Deoxyribose operon repressor.	<i>Escherichia coli</i>
o209#5	298779	298405	125	19.2	-	228	32.5	123 ( <i>bcrC</i> )		P42334	SP	BcrC protein.	<i>Bacillus licheniformis</i>
o209#6	299282	300511	410	19.2	+	339	22.4	362 ( <i>bcr</i> , <i>bicA</i> , <i>bicR</i> , <i>sur</i> , <i>suxA</i> )		P28246	SP	Bicyclomycin resistance protein (sulfonamide resistance protein).	<i>Escherichia coli</i>
o209#9	301743	300928	272	19.2	-	163	40.5	74		P21878	SP	Hypothetical protein in <i>pdhA</i> 5' region (ORF1) (fragment).	<i>Bacillus stearothermophilus</i>
o209#10	302948	301743	402	19.3	-	220	26	384 ( <i>mosC</i> )		Q07609	SP	Membrane protein MosC.	<i>Rhizobium meliloti</i>
o209#12	305428	303746	561	19.3	-	147	23.7	565 (HI0035)		P44472	SP	Hypothetical protein HI0035.	<i>Haemophilus influenzae</i>
o210#1	306362	306108	85	19.4	-	579	100	85 <i>grxA</i> , <i>grx</i>		P00277	SP	Glutaredoxin 1 (GRX1).	<i>Escherichia coli</i>
o210#2	306522	306806	95	19.4	+	600	100	95	<i>ybiC</i>	P46119	SP	Hypothetical 10.5 KD protein in <i>grxA-mdxA</i> intergenic region.	<i>Escherichia coli</i>
o210#3	306793	307512	240	19.4	+	1599	100	240 <i>mdaA</i> , <i>mda18</i>		P17117	SP	Modulator of drug activity A.	<i>Escherichia coli</i>
o210#4	307819	308475	219	19.4	+	1817	100	282 <i>rimK</i>		P17116	SP	Ribosomal protein S6 modification protein.	<i>Escherichia coli</i>
o210#6	309393	310502	370	19.4	+	2442	99.5	370 <i>potF</i>		A45313	PIR	Putrescine transport protein PotF	<i>Escherichia coli</i>

Table 1. Continued.

orf number	start* (nt)	end (nt)	length (aa)	map (min)	direction*	score*	%*	overlap*	gene'	y gene*	accession*	product	species	
o210#7	310519	311730	404	19.4	+	2688	100	404	<i>potG</i>	P31134	SP	Putrescine transport ATP-binding protein PotG.	<i>Escherichia coli</i>	
o210#8	311743	312693	317	19.5	+	2105	100	317	<i>potH</i>	P31135	SP	Putrescine transport system permease protein PotH.	<i>Escherichia coli</i>	
o210#9	312693	313535	281	19.5	+	1807	100	281	<i>potI</i>	P31136	SP	Putrescine transport system permease protein PotI.	<i>Escherichia coli</i>	
o211#1	314127	315251	375	19.5	+	1444	54.6	388	(HI0958)	P44083	SP	Hypothetical protein HI0958.	<i>Haemophilus influenzae</i>	
o211#2	316184	315456	243	19.6	-	1574	99.6	243	<i>artJ</i>	P30860	SP	Arginine-binding periplasmic protein 2 precursor.	<i>Escherichia coli</i>	
o211#3	317143	316478	222	19.6	-	1394	98.6	222	<i>artM</i>	P30862	SP	Arginine transport system permease protein ArtM.	<i>Escherichia coli</i>	
o211#4	317859	317146	238	19.6	-	1498	98.7	238	<i>artQ</i>	S31729	pir	Arginine transport system protein ArtQ	<i>Escherichia coli</i>	
o211#5	318597	317869	243	19.6	-	1531	97.1	244	<i>artI</i>	P30859	SP	Arginine-binding periplasmic protein 1 precursor.	<i>Escherichia coli</i>	
o211#6	319343	318618	242	19.6	-	1542	98.8	242	<i>artP</i>	P30858	SP	Arginine transport ATP-binding protein ArtP.	<i>Escherichia coli</i>	
o211#8	320202	320522	107	19.7	+	347	50.5	103		P33382	SP	Hypothetical 12.0 KD protein in <i>plcB-ldh</i> intergenic region (ORFB).	<i>Listeria monocytogenes</i>	
o211#9	320522	321349	276	19.7	+	173	41.1	107	( <i>ampD</i> )	Q00831	SP	AmpD protein.	<i>Citrobacter freundii</i> , and <i>Enterobacter cloacae</i>	
o211#10	322398	321352	349	19.7	-	147	25.3	356		P27364	SP	3-beta hydroxy-5-ene steroid dehydrogenase type III (EC 1.1.1.145) / steroid delta-isomerase (EC 5.3.3.1)	<i>Rattus norvegicus</i> (Rat)	
o211#12	324903	323905	333	19.7	-	647	36.9	331	( <i>GLY1</i> , YEL046C, SYGP-ORF34)	P37303	SP	GLY1 protein.	<i>Saccharomyces cerevisiae</i> (Baker's Yeast).	
o211#13	326658	324943	572	19.8	-	3776	100	572	<i>poxB</i>	A23648	pir	Pyruvate dehydrogenase (cytochrome) (EC 1.2.2.2)	<i>Escherichia coli</i>	
o211#14	327765	326794	324	19.8	-	195	42.3	78		P00235	SP	Ferredoxin 1.	<i>Equisetum arvense</i> (Field horsetail)	
o211#15	329423	327774	550	19.8	-	1375	40.5	561		P31101	SP	Prismane protein.	<i>Desulfovibrio vulgaris</i> (Strain Hildenborough).	
o211#16	330466	329570	299	19.9	-	638	47.4	190	(HI1452)	P44201	SP	Hypothetical protein HI1452.	<i>Haemophilus influenzae</i>	
o211#17	331656	330964	231	19.9	-	302	38.8	232		JQ2285	pir	Nodulin-26	Soybean	
o212#2	334729	333740	330	20.0	-	606	39.2	306	( <i>virK</i> )	S25264	pir	VirK protein	<i>Shigella flexneri</i> plasmid pMYSH6000	
o212#3	334844	335956	371	20.0	+	200	30.8	208	( <i>mtrC</i> )	S40252	pir	MtrC protein	<i>Neisseria gonorrhoeae</i>	
o212#4	335956	337899	648	20.0	+	572	39.5	223		S38423	pir	DevA protein	<i>Anabaena</i> sp.	
o212#5	338199	337978	74	20.0	-	508	100	74	<i>cspD</i> , <i>cspH</i>	P24245	SP	Cold shock-like protein CspD.	<i>Escherichia coli</i>	
o212#7	338873	341146	758	20.1	+	4831	99.6	758	<i>clpA</i> , <i>lopD</i>	P15716	SP	ATP-dependent Clp protease ATP-binding subunit ClpA.	<i>Escherichia coli</i>	
o212#8	342052	341837	72	20.1	-	457	100	72	<i>infA</i>	C36888	pir	Translation initiation factor eIF-1	<i>Escherichia coli</i>	
o212#9	342162	341983	60	20.1	-	420	100	60		JS0583	pir	Hypothetical 7K protein ( <i>infA</i> 5' region)	<i>Escherichia coli</i>	
o214#1	342737	342165	191	20.1	-	1264	99.5	191		JS0582	pir	Hypothetical 22K protein ( <i>infA</i> 5' region)	<i>Escherichia coli</i>	
o214#2	343041	342340	234	20.1	-	1656	100	234	<i>aat</i>	JS0581	pir	27K protein ( <i>infA</i> 5' region)	<i>Escherichia coli</i>	
o214#3	344804	343086	573	20.2	-	3643	100	573	<i>cydC</i> , <i>mdrA</i> , <i>mdrH</i> , <i>surB</i>	P23886	SP	Transport ATP-binding protein CydC.	<i>Escherichia coli</i>	
o214#4	346571	344808	588	20.2	-	3806	99.8	588	<i>cydD</i> , <i>htrD</i>	P29018	SP	Transport ATP-binding protein CydD.	<i>Escherichia coli</i>	
o214#5	347659	346697	321	20.2	-	2096	100	321	<i>trxB</i>	A28074	pir	Thioredoxin reductase (NADPH) (EC 1.6.4.5)	<i>Escherichia coli</i>	
o214#6	348204	348695	164	20.3	+	1061	100	164	<i>lrp</i> , <i>alsB</i> , <i>livR</i>	P19494	SP	Leucine-responsive regulatory protein.	<i>Escherichia coli</i> , and <i>Enterobacter aerogenes</i> ( <i>Aerobacter aerogenes</i> )	
o214#7	348833	352819	1329	20.3	+	8853	99.3	1329	<i>ftsK</i>	P46889	SP	Cell division protein FtsK.	<i>Escherichia coli</i>	
o214#8	352978	353589	204	20.4	+	1363	100	204	<i>lolA</i> , <i>lplA</i>	P39178	SP	Outer membrane lipoproteins carrier protein precursor (P20).	<i>Escherichia coli</i>	
o214#9	353603	354943	447	20.4	+	551	98.8	81		P45526	SP	Hypothetical protein in <i>serS</i> 5' region (fragment).	<i>Escherichia coli</i>	
o215#1	355037	356326	430	20.4	+	2836	100	430	<i>serS</i>	A26400	pir	Serine-tRNA ligase (EC 6.1.1.11)	<i>Escherichia coli</i>	
o215#2	356655	359009	785	20.4	+	5358	100	785	<i>dmsA</i>	S03785	pir	Dimethylsulfoxide reductase chain A	<i>Escherichia coli</i>	
o215#3	359023	359637	205	20.5	+	1489	99	207	<i>dmsB</i>	S03786	pir	Dimethylsulfoxide reductase (EC 1.8.-.-) chain B	<i>Escherichia coli</i>	
o215#4	359642	360502	287	20.5	+	1896	100	287	<i>dmsC</i>	S03787	pir	Dimethylsulfoxide reductase chain C	<i>Escherichia coli</i>	
o215#5	361166	360543	208	20.5	-	1383	100	208		ycuC	P21367	SP	Hypothetical 23.1 KD protein in <i>dmsC</i> 3' region.	<i>Escherichia coli</i>

Table 1. Continued.

orf number	start* (nt)	end (nt)	length (aa)	map (min)	direction*	score*	%*	overlap* gene'	y gene'	accession*	product	species	
o215#6	361480	362625	382	20.6	+	179	21.6	402 ( <i>gtr, glcP</i> )	<i>ycaD</i>	P15729	SP	Glucose transport protein.	<i>Synechocystis SP.</i> (Strain PCC 6803).
o216#1	362832	364265	478	20.6	+	476	23	426 ( <i>ygjI</i> )		P42590	SP	Hypothetical 52.1 KD protein in <i>ebgC-exuT</i> intergenic region (O477).	<i>Escherichia coli</i>
o216#2	365177	364272	302	20.6	-	648	42.7	248		Q01610	SP	Putative transcriptional regulator (fragment).	<i>Pseudomonas</i>
o216#3	365277	365864	196	20.6	+	427	100	64	<i>ycaK</i>	P43340	SP	Hypothetical protein in <i>pflA</i> 3' region (fragment).	<i>Escherichia coli</i>
o216#4	366686	365952	245	20.7	-	1696	100	245 <i>pflA, act</i>		P09374	SP	Pyruvate formate-lyase 1 activating enzyme (EC 1.97.1.4).	<i>Escherichia coli</i>
o216#5	369163	366884	760	20.7	-	5020	100	760 <i>pfl</i>		S01788	PIR	Formate C-acetyltransferase (EC 2.3.1.54)	<i>Escherichia coli</i>
o216#6	370075	369221	285	20.7	-	1885	99.6	285 <i>focA</i>		A32305	PIR	Probable formate transporter	<i>Escherichia coli</i>
o216#7	372250	370484	589	20.7	-	2656	66.9	583 (HI1265)		P44144	SP	Hypothetical protein HI1265.	<i>Haemophilus influenzae</i>
o216#9	373262	374347	362	20.8	+	2400	99.7	362 <i>serC</i>		S28806	PIR	Phosphoserine transaminase (EC 2.6.1.52)	<i>Escherichia coli</i>
o216#10	374421	375701	427	20.8	+	2760	99.5	427 <i>aroA</i>		A30370	PIR	3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19)	<i>Escherichia coli</i>
o216#11	375849	376634	262	20.9	+	944	56.1	253	<i>ycaL</i>	A42604	PIR	ORF1 3' to <i>speB</i>	<i>Escherichia coli</i>
o216#12	376810	377490	227	20.9	+	1425	99.6	227 <i>cmk, mssA</i>		P23863	SP	Cytidylate kinase (EC 2.7.4.14)	<i>Escherichia coli</i>
o217#1	377604	379274	557	20.9	+	3546	99.8	557 <i>rpsA, ssyF</i>		P02349	SP	30S ribosomal protein S1.	<i>Escherichia coli</i>
o217#2	379437	379718	94	20.9	+	616	100	94 <i>himD, hip</i>		P08756	SP	Integration host factor beta-subunit ( <i>ihf-beta</i> ).	<i>Escherichia coli</i>
o217#3	379851	382190	780	21.0	+	715	32.8	799 ( <i>rec2, rec-2, HI0061</i> )	<i>ycaI</i>	P44408	SP	Recombination protein 2.	<i>Haemophilus influenzae</i>
o217#4	382230	383975	582	21.0	+	3669	100	582 <i>msbA</i>		S27998	PIR	MsbA protein	<i>Escherichia coli</i>
o218#1	383975	384958	328	21.0	+	2202	99.7	328	<i>ycaH</i>	P27300	SP	Hypothetical 35.6 KD protein in <i>msbA-kdsB</i> intergenic region (ORFE).	<i>Escherichia coli</i>
o218#4	386461	387204	248	21.1	+	1637	100	248 <i>kdsB</i>		A26322	PIR	3-deoxy-manno-octulosonate cytidyltransferase (EC 2.7.7.38)	<i>Escherichia coli</i>
o218#6	389010	388234	259	21.1	-	971	89.4	170 ( <i>ycbC</i> )		P36565	SP	Hypothetical protein in <i>kdsB-kicB</i> intergenic region (fragment).	<i>Escherichia coli</i>
o218#7	389146	389928	261	21.2	+	1772	99.6	261 <i>stmA</i>		P36566	SP	Hypothetical 29.8 KD protein in <i>kdsB-kicB</i> intergenic region.	<i>Escherichia coli</i>
o218#8	389928	391247	440	21.2	+	2863	100	440 <i>mukF, kicB</i>		P36567	SP	Killing factor KicB.	<i>Escherichia coli</i>
o218#9	391204	391932	243	21.2	+	1589	100	243 <i>mukE, kicA</i>		S43912	PIR	KicA protein	<i>Escherichia coli</i>
o218#10	391935	396392	1486	21.2	+	8605	97.8	1391 <i>mukB</i>		P22523	SP	MukB protein	<i>Escherichia coli</i>
o218#11	396656	398500	615	21.3	+	396	96.8	63	<i>ycbB</i>	P22525	SP	Hypothetical protein in <i>mukB</i> 3' region (fragment).	<i>Escherichia coli</i>
o218#12	398684	399229	182	21.4	+	375	45.2	126 (HI1666)		P44284	SP	Hypothetical protein HI1666.	<i>Haemophilus influenzae</i>
o218#13	399259	399903	215	21.4	+	155	27.9	183		S45349	PIR	L1 metallo-beta-lactamase	<i>Xanthomonas maltophilia</i>
o218#14	401318	400131	396	21.4	-	2601	100	396 <i>aspC</i>		A00598	PIR	Aspartate transaminase (EC 2.6.1.1)	<i>Escherichia coli</i>
o218#15	402591	401506	362	21.4	-	2389	100	362 <i>ompF, toIF, cmlB, coa, cry</i>		P02931	SP	Outer membrane protein F precursor (Outer membrane protein 1A, IA, OR B).	
o219#1	404594	403197	466	21.5	-	3098	100	466 <i>asnS</i>		S29297	PIR	Asparagine-tRNA ligase (EC 6.1.1.22)	<i>Escherichia coli</i>
o219#2	405965	404766	400	21.5	-	2698	100	400 <i>pncB</i>		JQ0756	PIR	Nicotinate phosphoribosyltransferase (EC 2.4.2.11)	<i>Escherichia coli</i>
o219#3	406234	408840	869	21.5	+	5760	100	869 <i>pepN</i>		P04825	SP	Aminopeptidase N (EC 3.4.11.2)	<i>Escherichia coli</i>
o219#4	409653	408889	255	21.6	-	566	39.2	260 ( <i>nrtD</i> )	<i>ycbE</i>	P38046	SP	Nitrate transport ATP-binding protein NrtD.	<i>Synechococcus SP.</i> (Strain PCC 7942)
o221#1	410486	409653	278	21.6	-	367	29.9	274 ( <i>nrtB</i> )		P38044	SP	Nitrate transport permease protein NrtB.	<i>Synechococcus SP.</i> (Strain PCC 7942)
o221#2	411597	410455	381	21.6	-	1647	64.7	371 ( <i>yjeC</i> )		P40402	SP	Hypothetical protein (ORFM) (fragment).	<i>Bacillus subtilis</i>
o221#3	412559	411597	321	21.6	-	547	32.4	284 ( <i>yjeA</i> )		P40400	SP	Hypothetical 36.3 KD lipoprotein precursor (ORFK).	<i>Bacillus subtilis</i>
o221#5	413469	414014	182	21.7	+	290	32.7	171 (F17A)		P11312	SP	F17 fimbrial protein precursor (F17 Pilin)	<i>Escherichia coli</i>
o221#6	414100	414798	233	21.7	+	588	45.3	201 ( <i>focC</i> )		P46008	SP	Chaperone protein FocC precursor.	<i>Escherichia coli</i>
o221#7	414826	417423	866	21.7	+	2773	48.5	881 ( <i>fimD</i> )		P30130	SP	Outer membrane usher protein FimD precursor.	<i>Escherichia coli</i>
o221#8	417417	418484	356	21.8	+	229	22.8	355 ( <i>yraK</i> )		P43319	SP	Hypothetical 38.5 KD protein in <i>agal-mtr</i> intergenic region.	<i>Escherichia coli</i>
o221#9	418499	419038	180	21.8	+	269	30.5	187		B28393	PIR	Type 1 fimbrial protein precursor	<i>Salmonella typhimurium</i>
o221#10	419001	419561	187	21.8	+	297	38.5	156 ( <i>fimF</i> )		P08189	SP	FimF protein precursor.	<i>Escherichia coli</i>
o221#11	419530	420264	245	21.8	+	413	32.3	223 ( <i>fimC</i> )	<i>ycbF</i>	P31697	SP	Chaperone protein FimC precursor.	<i>Escherichia coli</i>

Table 1. Continued.

orf number	start* (nt)	end (nt)	length (aa)	map (min)	direction <sup>a</sup>	score <sup>c</sup>	% <sup>d</sup>	overlap <sup>e</sup>	gene <sup>f</sup>	y gene <sup>g</sup>	accession <sup>h</sup>	product	species
o221#12	420378	421385	336	21.8	+	2206	100	336	<i>pyrD</i>		A23109	Pir Dihydroorotate oxidase (EC 1.3.3.1)	<i>Escherichia coli</i>
o221#14	423168	422104	355	21.9	-	193	42.6	68	( <i>rfbI</i> )		P26395	SP RfbI protein.	<i>Salmonella typhimurium</i>
o221#15	423454	425559	702	21.9	+	1495	60.7	529	(HI0115)		P44524	SP Hypothetical protein HI0115.	<i>Haemophilus influenzae</i>
o222#1	425574	427478	635	21.9	+	691	30.3	531	(GCN20, YFR009W)		P43535	SP GCN20 protein.	<i>Saccharomyces cerevisiae</i>
o222#2	427611	428861	417	22.0	+	2422	94.6	390	<i>pqiA</i>		P43670	SP Paraquat-inducible protein A.	<i>Escherichia coli</i>
o222#3	428869	430506	546	22.0	+	595	100	92	<i>pqiB</i> , <i>pqi5B</i>		P43671	SP Paraquat-inducible protein B (fragment).	<i>Escherichia coli</i>
o222#5	431325	431489	55	22.1	+	383	98.2	55	<i>rmf</i>		P22986	SP Ribosome modulation factor (protein E). S36191:	<i>Escherichia coli</i>
o222#6	431565	432080	172	22.1	-	1156	100	169	<i>fabA</i>		P18391	SP 3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase (EC	<i>Escherichia coli</i>
o222#7	433318	432152	389	22.1	-	1339	41.9	570	( <i>lon-B</i> , HI1324)		P43865	SP Protease La homolog (EC 3.4.21.-).	<i>Haemophilus influenzae</i>
o222#8	434095	434544	150	22.1	+	233	100	39		<i>ycbG</i>	P45569	SP Hypothetical protein in <i>ompA</i> 3' region (fragment).	<i>Escherichia coli</i>
o222#9	435663	434626	346	22.1	-	2316	100	346	<i>ompA</i> , <i>tolG</i> , <i>tut</i> , <i>con</i>		P02934	SP Outer membrane protein A precursor (outer membrane protein II*).	<i>Escherichia coli</i>
o222#10	436529	436023	169	22.2	-	1095	100	169	<i>suaA</i> , <i>sfIA</i>		P08846	SP Cell division inhibitor.	<i>Escherichia coli</i>
o222#12	439502	437343	720	22.2	-	1541	38.5	702	(HI1680)		P44289	SP Hypothetical protein HI1680.	<i>Haemophilus influenzae</i>
o222#13	439958	439515	148	22.3	-	211	100	35		<i>yccF</i>	P37065	SP Hypothetical protein in <i>helD</i> 5' region (fragment).	<i>Escherichia coli</i>
o223#1	440081	442132	684	22.3	+	4501	99	684	<i>helD</i>		JV0021	pir Helicase (EC 3.6.1.-) IV	<i>Escherichia coli</i>
o223#2	442634	442170	155	22.3	-	1008	98	153		<i>yccG</i>	P37066	SP Hypothetical protein in <i>helD</i> 3' region (fragment).	<i>Escherichia coli</i>
o223#3	443383	442724	220	22.3	-	163	34.8	89			P29740	SP Hypothetical protein in <i>ompH</i> 3' region (ORF2) (fragment).	<i>Photobacterium sp.</i> (Strain SS9).
o223#6	445579	444392	396	22.4	-	659	32.9	392	( <i>ywbD</i> , <i>ipa-19D</i> )		P39587	SP Hypothetical 44.4 KD protein in <i>epr-galK</i> intergenic region.	<i>Bacillus subtilis</i>
o223#7	445461	445949	163	22.4	+	213	44.7	85			P07032	SP Acylphosphatase, organ-common type isozyme (EC 3.6.1.7)	<i>Gallus gallus</i> (Chicken).
o223#8	446278	445952	109	22.4	-	347	96.4	55		<i>yccK</i>	P45572	SP Hypothetical protein in <i>helD-serT</i> intergenic region (fragment).	<i>Escherichia coli</i>
o223#9	447028	446372	219	22.4	-	1351	100	219		<i>yccA</i>	P06967	SP Hypothetical 23.4 KD protein in <i>serT</i> 5' region.	<i>Escherichia coli</i>
o224#1	447749	448864	372	22.4	+	2585	100	372	<i>hyaA</i>		JV0072	pir Hydrogenase (EC 1.18.99.1) (NiFe) 1 small chain precursor	<i>Escherichia coli</i>
o224#2	448864	450654	597	22.5	+	4124	100	597	<i>hyaB</i>		P19927	SP Hydrogenase-1 large chain (EC 1.18.99.1)	<i>Escherichia coli</i>
o224#3	450676	451380	235	22.5	+	1635	100	235	<i>hyaC</i>		JV0074	pir HyaC protein	<i>Escherichia coli</i>
o224#4	451380	451964	195	22.5	+	1295	100	195	<i>hyaD</i>		P19930	SP Hydrogenase-1 operon protein HyaD.	<i>Escherichia coli</i>
o224#5	451964	452359	132	22.5	+	909	100	132	<i>hyaE</i>		P19931	SP Hydrogenase-1 operon protein HyaE.	<i>Escherichia coli</i>
o224#6	452359	453213	285	22.5	+	1936	100	285	<i>hyaF</i>		P19932	SP Hydrogenase-1 operon protein HyaF.	<i>Escherichia coli</i>
o224#7	453350	454891	514	22.6	+	3452	100	514	<i>appC</i>		S17958	pir AppC protein	<i>Escherichia coli</i>
o224#8	454906	456039	378	22.6	+	2572	99.7	378	<i>appB</i>		S17959	pir AppB protein	<i>Escherichia coli</i>
o224#9	456024	456194	57	22.6	+	246	97.4	38			C26534	pir Hypothetical protein 2 ( <i>appA</i> 5' region)	<i>Escherichia coli</i>
o224#10	456227	457522	432	22.6	+	2883	100	432	<i>appA</i>		P07102	SP Periplasmic phosphoanhydride phosphohydrolase precursor (EC 3.1.3.2) (ph 2.5 acid phosphatase) (ap) (6-phytase) (EC 3.1.3.26).	<i>Escherichia coli</i>
o225#1	459820	457643	726	22.6	-	432	32.2	227	( <i>capB</i> )	<i>yccC</i>	P39851	SP CapB protein.	<i>Staphylococcus aureus</i>
o225#4	463555	461462	698	22.7	-	3154	64.6	698	( <i>yjbH</i> )		P32689	SP Hypothetical 78.5 KD protein in <i>pgi-xylE</i> intergenic region (O698).	<i>Escherichia coli</i>
o225#5	463818	463558	87	22.8	-	671	42.2	244	( <i>yjbG</i> )		P32688	SP Hypothetical 26.3 KD protein in <i>pgi-xylE</i> intergenic region precursor (O245).	<i>Escherichia coli</i>
o225#6	464942	464301	214	22.8	-	539	38.7	212	( <i>yjbF</i> )		P32687	SP Hypothetical 25.0 KD lipoprotein in <i>pgi-xylE</i> intergenic region precursor (O222).	<i>Escherichia coli</i>
o225#8	465372	465052	107	22.8	-	129	38.4	86			Q05049	SP Integumentary mucin C.1 (FIM-C.1) (fragment).	<i>Xenopus laevis</i> (African clawed frog).
o225#9	465414	465623	70	22.8	+	462	98.6	70	IS1		S40546	pir Hypothetical protein	<i>Escherichia coli</i>
o225#10	465443	465715	91	22.8	+	620	98.9	91	<i>insA</i>		JN0138	pir Hypothetical protein InsA (insertion sequence IS1D)	<i>Escherichia coli</i>
o225#11	465637	466137	167	22.8	+	1160	100	167	<i>insB</i>		JN0139	pir Hypothetical protein InsB' (insertion sequence IS1D)	<i>Escherichia coli</i>
o225#12	466106	465819	96	22.8	-	638	99	96	IS1		D93826	pir Hypothetical 11K protein (insertion sequence IS1)	<i>Escherichia coli</i>
o225#13	466773	466576	66	22.8	-	357	75.7	70	( <i>cspF</i> )		P39819	SP Cold shock-like protein CspF.	<i>Escherichia coli</i>
o225#14	467071	467280	70	22.8	+	371	79.4	68	( <i>cspB</i> )		P36995	SP Cold shock-like protein CspB.	<i>Escherichia coli</i>

Table 1. Continued.

orf number	start* (nt)	end (nt)	length (aa)	map (min)	direction*	score*	%	overlap* gene*	y gene*	accession*	product	species
o226#1	468972	467902	357	22.9	-	277	29.3	188 ( <i>nosR</i> )		Q00790	SP Regulatory protein NosR.	<i>Pseudomonas stutzeri</i> ( <i>Pseudomonas perfectomarina</i> ).
o226#2	471788	469047	914	22.9	-	758	31.4	631 ( <i>lemA</i> )	<i>ycfI</i>	B41863	PIR Two-component regulatory protein LemA	<i>Pseudomonas syringae</i>
o226#3	471871	472896	342	23.0	+	2226	98.8	342	<i>ycfH</i>	P38683	SP Hypothetical 37.8 KD protein in <i>torR</i> 5' region precursor (ORF1).	<i>Escherichia coli</i>
o226#4	473564	472875	230	23.0	-	1473	99.6	230 <i>torR</i>		P38684	SP <i>torCAD</i> operon transcriptional regulatory protein TorR.	<i>Escherichia coli</i>
o226#5	473694	474863	390	23.0	+	2639	99.5	390 <i>torC</i>		P33226	SP Cytochrome C-type protein TorC.	<i>Escherichia coli</i>
o226#6	474866	477409	848	23.0	+	5726	98.5	848 <i>torA</i>		P33225	SP Trimethylamine-N-oxide reductase precursor (EC 1.6.6.9)	<i>Escherichia coli</i>
o226#7	477409	478005	199	23.1	+	1298	99.5	199 <i>torD</i>		P36662	SP TorD protein.	<i>Escherichia coli</i>
o226#8	478465	478163	101	23.1	-	657	100	101	<i>ycfD</i>	P36660	SP Hypothetical 11.5 protein in <i>torD-cbpA</i> intergenic region (ORF-2).	<i>Escherichia coli</i>
o226#9	479385	478468	306	23.1	-	1980	99	297 <i>cbpA</i>		P36659	SP Curved DNA-binding protein.	<i>Escherichia coli</i>
o227#1	479646	480899	418	23.1	+	1628	100	242	<i>ycfE</i>	P36661	SP Hypothetical protein in <i>cbpA-agp</i> intergenic region (ORF-D fragment).	<i>Escherichia coli</i>
o227#2	481195	482433	413	23.2	+	2753	100	413 <i>agp</i>		P19926	SP Glucose-1-phosphatase precursor (EC 3.1.3.10) (G1Pase).	<i>Escherichia coli</i>
o227#3	482701	482477	75	23.2	-	496	100	75	<i>ycfJ</i>	P46131	SP Hypothetical 8.5 KD protein in <i>agp</i> 3' region.	<i>Escherichia coli</i>
o227#4	483315	482725	197	23.2	-	1273	99.5	197 <i>wrbA</i>		P30849	SP Trp repressor binding protein.	<i>Escherichia coli</i>
o227#5	483661	483861	67	23.2	+	326	86.8	53 ( <i>yciG</i> )		P21361	SP Hypothetical 6.0 KD protein in <i>tonB-trpA</i> intergenic region (ORF1).	<i>Escherichia coli</i>
o227#6	485449	484124	442	23.2	-	960	39.8	422 ( <i>pyrP</i> )		P41006	SP Uracil permease ( <i>uracil transporter</i> ).	<i>Bacillus caldolyticus</i>
o227#7	486045	485473	191	23.2	-	414	42.2	161		C55349	PIR 4-hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.3) small chain	<i>Escherichia coli</i> (ATCC 11105)
o227#9	487375	486578	266	23.3	-	188	24.9	257 ( <i>bchO</i> )		P26174	SP Magnesium-chelataase 30 KD subunit.	<i>Rhodobacter capsulatus</i> ( <i>Rhodopseudomonas capsulata</i> )
o227#10	487769	487386	128	23.3	-	138	25.7	74		S30349	PIR Perchloric acid-soluble protein, 23K	<i>Rattus norvegicus</i> (Rat)
o228#1	488476	487784	231	23.3	-	158	25.1	175 ( <i>entB</i> , <i>entG</i> )		P15048	SP Isochorismatase (EC 3.3.2.1) (2,3-Dihydro-2,3-dihydroxybenzoate synthase).	<i>Escherichia coli</i>
o228#3	489852	490487	212	23.3	+	131	37.5	64 ( <i>mtrR</i> )		P39897	SP Regulatory protein MtrR.	<i>Neisseria gonorrhoeae</i>
o228#4	494492	490533	1320	23.4	-	8565	99.9	1320 <i>putA</i> , <i>poaA</i>		P09546	SP Proline dehydrogenase (EC 1.5.99.8) (proline oxidase) / delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12)	
o228#5	494915	496420	502	23.5	+	3279	100	502 <i>putP</i>		A30258	PIR Proline carrier protein	<i>Escherichia coli</i>
o228#6	497064	497792	243	23.5	+	203	25.6	227 ( <i>ywbL</i> , <i>ipa-27D</i> )		P39595	SP Hypothetical 52.4 KD protein in <i>epr-galk</i> intergenic region precursor.	<i>Bacillus subtilis</i>
o228#7	497892	498977	362	23.5	+	653	44.6	242 ( <i>ywbM</i> , <i>ipa-28D</i> )		P39596	SP Hypothetical 42.8 KD protein in <i>epr-galk</i> intergenic region.	<i>Bacillus subtilis</i> .
o229#1	498986	500254	423	23.5	+	1539	100	228	<i>ycfB</i>	P31545	SP Hypothetical protein in <i>phoH</i> 5' region (ORF1) (fragment).	<i>Escherichia coli</i>
o229#2	500602	501663	354	23.6	+	2334	100	354 <i>phoH</i> , <i>psiH</i>		P31544	SP PhoH protein (phosphate starvation-inducible protein PsiH).	<i>Escherichia coli</i>
o229#4	503456	502134	441	23.6	-	333	26.8	269		S18962	PIR Fbf15 protein	<i>Stigmatella aurantiaca</i>
o229#5	505466	503928	513	23.7	-	182	23.7	232 ( <i>yadE</i> )		P31666	SP Hypothetical 27.6 KD protein in <i>hpt-panD</i> intergenic region.	<i>Escherichia coli</i>
o229#7	508413	509840	476	23.7	+	172	27.5	160		P36892	SP Hypothetical 19.3 KD protein in <i>repSA</i> 5' region (ORF183).	<i>Streptomyces ambifaciens</i>
o229#8	510750	509887	288	23.8	-	2006	99.7	288 <i>IS3</i>		P05822	SP Putative transposase for insertion sequence IS3.	<i>Escherichia coli</i>
o230#3	513314	514249	312	23.9	+	281	31.4	204 ( <i>serA</i> )		P35136	SP D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (PGDH).	<i>Bacillus subtilis</i> .
o230#5	515068	515619	184	23.9	+	360	39.2	186 (HI1543)		P44248	SP Hypothetical protein HI1543.	<i>Haemophilus influenzae</i>
o231#1	518624	517977	216	24.0	-	175	56.9	58 ( <i>moaR</i> )		A36937	PIR Monoamine regulon positive regulator MoaR	<i>Klebsiella pneumoniae</i>
o231#2	519379	519831	151	24.0	+	114	100	18 <i>csxB</i>		P39828	SP Minor curlin subunit precursor (fragment).	<i>Escherichia coli</i>
o231#3	519875	520327	151	24.0	+	988	99.3	151 <i>csxA</i>		S31202	PIR Curlin precursor	<i>Escherichia coli</i>
o231#4	521725	520823	301	24.0	-	2068	99.7	301 <i>IS2</i>		JQ0040	PIR Hypothetical 34K protein (insertion sequence IS2)	<i>Escherichia coli</i>
o231#5	521249	521677	143	24.0	+	956	99.3	143 <i>IS2</i>		JQ0042	PIR Hypothetical 16K protein (insertion sequence IS2)	<i>Escherichia coli</i>

Table 1. Continued.

orf number	start* (nt)	end (nt)	length (aa)	map (min)	direction <sup>b</sup>	score <sup>c</sup>	% <sup>d</sup>	overlap <sup>e</sup>	gene <sup>f</sup>	y gene <sup>f</sup>	accession <sup>h</sup>	product	species	
o231#6	521652	521362	97	24.0	-	688	99	97	IS2		JQ0041	pir	Hypothetical 11K protein (insertion sequence IS2)	<i>Escherichia coli</i>
o231#7	522048	521686	121	24.0	-	773	100	121	IS2		JQ0039	pir	Hypothetical 13K protein (insertion sequence IS2)	<i>Escherichia coli</i>
o231#9	522584	523114	177	24.1	+	322	31.3	166			Q02874	SP	Histone macro-H2A.1.	<i>Rattus norvegicus</i> (Rat)
o231#10	523059	524537	493	24.1	+	187	23.2	444	( <i>cls, nov</i> )		P31071	SP	Cardiolipin synthetase (EC 2.7.8.-).	<i>Escherichia coli</i>
o232#1	526099	527631	511	24.1	+	3454	100	511	<i>mdoG</i>		S35417	pir	MdoG protein	<i>Escherichia coli</i>
o232#2	527627	530167	847	24.2	+	5776	99.8	847	<i>mdoH</i>		S35418	pir	MdoH protein	<i>Escherichia coli</i>
o232#3	530367	530567	67	24.2	+	474	100	67		<i>yceK</i>	P45806	SP	Hypothetical 7.5 KD protein in <i>mdoH-msyB</i> intergenic region.	<i>Escherichia coli</i>
o232#4	530945	530574	124	24.2	-	853	100	124	<i>msyB</i>		P25738	SP	Acidic protein MsyB.	<i>Escherichia coli</i>
o232#5	532254	531031	408	24.2	-	313	24.6	362	( <i>tetA</i> )	<i>yceE</i>	P02981	SP	Tetracycline resistance protein.	<i>Escherichia coli</i>
o232#6	533346	532429	306	24.3	-	2152	100	306	<i>htrB</i>		P24187	SP	Heat shock protein B.	<i>Escherichia coli</i>
o232#7	533571	534620	350	24.3	+	2410	99.1	350		<i>yceA</i>	P24188	SP	Hypothetical 40.0 KD protein in <i>htrB</i> 5' region (ORF39.9).	<i>Escherichia coli</i>
o232#8	535240	534668	191	24.3	-	1243	100	191	<i>yceI</i>		P37904	SP	18.7 KD protein in <i>htrB</i> 5' region precursor.	<i>Escherichia coli</i>
o232#9	535810	535247	188	24.3	-	201	27.6	163	( <i>cybB</i> )		P08732	SP	Cytochrome B561.	<i>Escherichia coli</i>
o232#10	537350	536235	372	24.4	-	972	41.4	379	( <i>soxA</i> )		P40873	SP	Sarcosine oxidase (EC 1.5.3.1).	<i>Arthrobacter sp.</i> (Strain TE1826).
o232#13	539371	538328	348	24.4	-	2322	100	348	<i>pyrC</i>		A25008	pir	Dihydroorotase (EC 3.5.2.3)	<i>Escherichia coli</i>
o233#1	540037	539480	186	24.4	-	1182	100	186		<i>yceB</i>	P09995	SP	Hypothetical 20.5 KD protein in <i>pyrC</i> 3' region.	<i>Escherichia coli</i>
o233#3	542120	540885	412	24.5	-	233	20.6	399	( <i>blt, bmt, bmr2</i> )		P39843	SP	Multidrug resistance protein 2 (Multidrug-efflux transporter 2).	<i>Bacillus subtilis</i> .
o233#4	542326	542907	194	24.5	+	1370	100	194	<i>rimJ</i>		P09454	SP	Ribosomal-protein-alanine acetyltransferase (EC 2.3.1.128)	<i>Escherichia coli</i>
o233#5	542921	543565	215	24.5	+	1236	93.9	212	( <i>yceH</i> )		P29217	SP	Hypothetical 25.6 KD protein in <i>rimJ</i> 3' region (G20.3).	<i>Escherichia coli</i>
o233#6	543570	544490	307	24.5	+	1709	80.8	307	( <i>mviM</i> )		P37168	SP	Virulence factor MviM.	<i>Salmonella typhimurium</i>
o233#7	544603	546135	511	24.5	+	3118	92.7	523	( <i>mviN</i> )		P37169	SP	Virulence factor MviN.	<i>Salmonella typhimurium</i>
o233#8	546594	546181	138	24.6	-	239	100	38	<i>flgN</i>		P43533	SP	Flagella synthesis protein FlgN (fragment).	<i>Escherichia coli</i>
o233#9	546892	546602	97	24.6	-	589	100	97	<i>flgM</i>		P43532	SP	Negative regulator of flagellin synthesis (anti-sigma factor).	<i>Escherichia coli</i>
o233#10	547627	546971	219	24.6	-	1059	71.2	219	( <i>flgA</i> )		P40131	SP	Flagella basal body P-ring formation protein FlgA precursor.	<i>Salmonella typhimurium</i>
o233#11	547782	548195	138	24.6	+	745	81.2	138	( <i>flgB</i> )		P16437	SP	Putative flagellar basal-body rod protein FlgB	<i>Salmonella typhimurium</i>
o233#12	548202	548603	134	24.6	+	794	93.3	134	( <i>flgC, flaW, fla</i> )		P16438	SP	Putative flagellar basal-body rod protein FlgC	<i>Salmonella typhimurium</i>
o233#14	548618	549310	231	24.6	+	1250	84.1	232	( <i>flgD</i> )		P16321	SP	Basal-body rod modification protein FlgD	<i>Salmonella typhimurium</i>
o233-234#1	549338	550543	402	24.6	+	2268	88.1	404	( <i>flgE</i> )		P16322	SP	Flagellar hook protein	<i>Salmonella typhimurium</i>
o234#2	550566	551318	251	24.7	+	1424	86.9	251	( <i>flgF, flaX, fla</i> )		P16323	SP	Putative flagellar basal-body rod protein FlgF	<i>Salmonella typhimurium</i>
o234#3	551493	552272	260	24.7	+	1636	99.2	260	( <i>flgG, flaL, fla</i> )		P16439	SP	Flagellar basal-body rod protein FlgG (distal rod protein).	<i>Salmonella typhimurium</i>
o234#4	552319	553023	235	24.7	+	1420	90.9	232	( <i>flgH, flaY, fla</i> )		P15929	SP	Flagellar L-ring protein precursor (basal body L-ring protein).	<i>Salmonella typhimurium</i>
o234#5	553038	554132	365	24.7	+	2135	92.1	365	( <i>flgI, flaM, fla</i> )		P15930	SP	Flagellar P-ring protein precursor.	<i>Salmonella typhimurium</i>
o234#6	554135	555073	313	24.7	+	1697	83.2	316	( <i>flgJ, flaZ, fla</i> )		P15931	SP	Flagellar protein FlgJ. D32887.	<i>Salmonella typhimurium</i>
o234#7	555142	556782	547	24.8	+	72	100	12	( <i>flgK, flaS, flaW</i> )		P33235	SP	Flagellar hook-associated protein 1 (Hap1) (fragment).	<i>Escherichia coli</i>
o234#8	556797	557747	317	24.8	+	1972	100	317	( <i>flgL, flaT, flaU</i> )		P29744	SP	Flagellar hook-associated protein 3 (Hap3) (hook-filament junction protein).	<i>Escherichia coli</i>
o234#9	561131	557949	1061	24.8	-	6861	99.6	1061	( <i>rne, ams, hmp1</i> )		P21513	SP	Ribonuclease E (EC 3.1.4.-) (RNase E).	<i>Escherichia coli</i>
o234#11	561704	562660	319	24.9	+	310	26.6	290		<i>yceC</i>	S50972	pir	DraP deaminase	<i>Saccharomyces cerevisiae</i>
o234#12	563398	562778	207	24.9	-	354	35.2	182	( <i>maf</i> )	<i>yceF</i>	Q02169	SP	Maf protein.	<i>Bacillus subtilis</i> .
o234#13	563558	564076	173	24.9	+	1139	100	173	( <i>g30K</i> )	<i>yceD</i>	P14189	SP	Hypothetical 19.3 KD protein in <i>rne-rpmF</i> intergenic region (G30K).	<i>Escherichia coli</i>
o234#14	564131	564301	57	25.0	+	368	100	57	<i>rpmF</i>		JV0048	pir	Ribosomal protein L32	<i>Escherichia coli</i>
o235#1	564415	565452	346	25.0	+	2197	100	346	<i>plsX</i>		P27247	SP	PlsX protein.	<i>Escherichia coli</i>
o235#2	565523	566473	317	25.0	+	2042	100	317	<i>fabH</i>		A42431	pir	3-oxoacyl-[acyl-carrier-protein] synthase (EC 2.3.1.41) III	<i>Escherichia coli</i>
o236#1	566492	567418	309	25.0	+	2003	100	309	<i>fabD</i>		B41856	pir	[acyl-carrier-protein] S-malonyltransferase (EC 2.3.1.39)	<i>Escherichia coli</i>
o236#2	567434	568165	244	25.0	+	1512	99.6	244	<i>fabG</i>		B42147	pir	3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100)	<i>Escherichia coli</i>
o236#3	568379	568612	78	25.1	+	477	100	78	<i>acpP</i>		C42147	pir	Acyl carrier protein	<i>Escherichia coli</i>



Table 1. Continued.

orf number	start* (nt)	end (nt)	length (aa)	map (min)	direction <sup>a</sup>	score <sup>c</sup>	% <sup>d</sup>	overlap <sup>e</sup>	gene <sup>f</sup>	y gene <sup>g</sup>	accession <sup>h</sup>	product	species
o236#4	568706	569941	412	25.1	+	2676	100	412	<i>fabF, fabI</i>		P39435	SP 3-oxoacyl-[acyl-carrier-protein] synthase II (EC 2.3.1.41)	<i>Escherichia coli</i>
o236#5	570064	570870	269	25.1	+	1794	100	269	<i>pabC</i>		P28305	SP 4-amino-4-deoxychorismate lyase (EC 4.-.-.-) (ADC LYASE).	<i>Escherichia coli</i>
o236#6	570876	571748	291	25.1	+	1579	100	243		<i>yceG</i>	P28306	SP Hypothetical protein in <i>pabC-holB</i> intergenic region (fragment).	<i>Escherichia coli</i>
o236#7	571691	571894	68	25.1	+	201	54.5	55	(HI0457)		P44720	SP Hypothetical protein HI0457.	<i>Haemophilus influenzae</i>
o236#8	571887	572525	213	25.1	+	770	96.7	121		<i>ycfG</i>	P37345	SP Hypothetical protein in <i>pabC-holB</i> intergenic region (fragment).	<i>Escherichia coli</i>
o236#9	572525	573526	334	25.1	+	2261	100	334	<i>holB</i>		B46738	PIR DNA-directed DNA polymerase (EC 2.7.7.7) III delta' chain	<i>Escherichia coli</i>
o236#10	573540	574334	265	25.2	+	701	99	104		<i>ycfH</i>	P37346	SP Hypothetical protein in <i>holB</i> 3' region (fragment).	<i>Escherichia coli</i>
o236#11	574632	576062	477	25.2	+	3104	100	477	<i>ptsG, glcA, umg</i>		P05053	SP PTS system.glucose- permease IIBC component (EC 2.7.1.69)	<i>Escherichia coli</i>
o236#12	578314	576128	729	25.2	-	4852	99.9	729	<i>fhuE</i>		S09262	PIR Outer membrane protein FhuE precursor	<i>Escherichia coli</i>
o236#13	578648	579004	119	25.3	+	383	55.6	108	( <i>zbp14, pkcI</i> )		P42856	SP 14 KD Zinc-binding protein.	<i>Zea mays</i> (maize)
o236#15	579398	580039	214	25.3	+	144	41.7	48			S19702	PIR Fibronectin-binding protein B	<i>Staphylococcus aureus</i>
o236#17	580858	581880	341	25.3	+	1232	51.7	344	(HI0959)		P44955	SP Hypothetical protein HI0959.	<i>Haemophilus influenzae</i>
o237#1	582848	584149	434	25.4	+	2833	99.8	434	<i>ndh</i>		A00461	PIR NADH dehydrogenase (EC 1.6.99.3)	<i>Escherichia coli</i>
o237#2	584362	584898	179	25.4	+	202	33.9	109	( <i>pcp</i> )	<i>ycfJ</i>	P31484	SP Outer membrane lipoprotein <i>pcp</i> precursor.	<i>Yersinia enterocolitica</i>
o237#4	585836	586090	85	25.4	+	230	46.5	86	( <i>ybiJ</i> )		P41038	SP Hypothetical 8.6 KD protein in <i>dinG/rarB</i> 3' region.	<i>Escherichia coli</i>
o237#5	587137	586178	320	25.4	-	940	48.2	307	( <i>ycfG</i> )		P40680	SP 31.6 KD protein in CobT 3' region precursor.	<i>Salmonella typhimurium</i>
o237#6	590727	587284	1148	25.5	-	7529	99.9	1148	<i>mfd</i>		A46215	PIR Transcription-repair coupling protein Mfd	<i>Escherichia coli</i>
o238#1	592139	593386	416	25.6	+	1147	46.1	401	(HI1555)		P44252	SP Hypothetical protein HI1555.	<i>Haemophilus influenzae</i>
o238#2	593292	594080	263	25.6	+	530	40.9	215	( <i>devA</i> )		S38423	PIR DevA protein	<i>Anabaena</i> sp.
o238#3	594083	595324	414	25.6	+	1439	51.3	417	(HI1548)		P44250	SP Hypothetical protein HI1548.	<i>Haemophilus influenzae</i>
o238#4	595356	596264	303	25.6	+	605	39	259	( <i>araJ</i> )		S27550	PIR AraJ protein	<i>Escherichia coli</i>
o238#8	599592	598549	348	25.7	-	2326	100	348	<i>potD</i>		P23861	SP Spermidine/putrescine-binding periplasmic protein precursor (SpbP).	<i>Escherichia coli</i>
o238#9	600383	599592	264	25.7	-	1665	100	264	<i>potC</i>		P23859	SP Spermidine/putrescineE transport system permease protein PotC.	<i>Escherichia coli</i>
o238#10	601207	600383	275	25.7	-	1772	99.6	275	<i>potB</i>		P23860	SP Spermidine/putrescine transport system permease protein PotB.	<i>Escherichia coli</i>
o239#1	602357	601224	378	25.8	-	2466	100	378	<i>potA</i>		P23858	SP Spermidine/putrescine transport ATP-binding protein PotA.	<i>Escherichia coli</i>
o239#2	602607	603830	408	25.8	+	271	100	42	<i>pepT</i>		P29745	SP Peptidase T (EC 3.4.11.-) (aminotripeptidase) (tripeptidase) (fragment).	<i>Escherichia coli</i>
o239#3	605003	603885	373	25.8	-	1998	100	280		<i>ycfD</i>	P27431	SP Hypothetical protein in <i>pepT-phoQ</i> intergenic region (fragment).	<i>Escherichia coli</i>
o239#4	606539	605082	486	25.8	-	3144	100	486	<i>phoQ</i>		B41966	PIR Virulence membrane protein PhoQ	<i>Escherichia coli</i>
o239#5	607210	606542	223	25.9	-	1434	100	223	<i>phoP</i>		A41965	PIR Alkaline phosphatase regulatory protein	<i>Escherichia coli</i>
o239#6	608749	607382	456	25.9	-	2979	100	456	<i>purB</i>		P25739	SP Adenylosuccinate lyase (EC 4.3.2.2)	<i>Escherichia coli</i>
o239#7	609394	608756	213	25.9	-	1335	99.5	213		<i>ycfC</i>	P25746	SP Hypothetical 22.9 KD protein in <i>purB</i> 5' region (ORF-23).	<i>Escherichia coli</i>
o239#8	610536	609433	368	25.9	-	1556	100	231		<i>ycfB</i>	P25745	SP Hypothetical protein in <i>purB</i> 5' region (ORF-15) (fragment).	<i>Escherichia coli</i>
o239#10	611714	611064	217	26.0	-	767	56.5	223	(HI0694)		P44827	SP Hypothetical protein HI0694.	<i>Haemophilus influenzae</i>
o239#11	611886	613133	416	26.0	+	2761	100	416	<i>icdA, icd, icdE</i>		P08200	SP Isocitrate dehydrogenase (NADP) (EC 1.1.1.42)	<i>Escherichia coli</i>
o240#1	615458	616348	297	26.1	+	1854	95.6	297	<i>lit</i>		A30386	PIR Lit protein	<i>Escherichia coli</i>
o240#2	617569	616445	375	26.1	-	1542	59.9	372	( <i>int</i> )		P27077	SP Integrase.	Bacteriophage P21
o240#3	617795	617553	81	26.1	-	203	36.4	77	( <i>xis</i> )		P27079	SP Excisionase.	Bacteriophage P21
o240#4	619696	619025	224	26.2	-	286	41.1	124	(C2)		P03035	SP Repressor protein C2.	Bacteriophage P21
o240#11	622855	623682	276	26.2	+	380	32.2	258	(HI1520)		P44240	SP Hypothetical protein HI1520.	<i>Haemophilus influenzae</i>
o240#13	624264	624890	209	26.3	+	744	99.1	114		<i>ycfK</i>	P45581	SP Hypothetical 13.1 KD protein in <i>pin</i> 5' region.	<i>Escherichia coli</i>
o240#14	625284	625003	94	26.3	-	631	98.9	94		<i>ycfA</i>	P09153	SP Hypothetical 32.8 KD protein in <i>pin</i> 5' region.	<i>Escherichia coli</i>
o240#15	625024	625284	87	26.3	+	573	98.9	87			P09154	SP Very Hypothetical 16.7 KD protein in invertible-P region of excisable element e14.	<i>Escherichia coli</i>
o240#16	625286	625459	58	26.3	+	420	100	58			P09154	SP Very Hypothetical 16.7 KD protein in invertible-p region of excisable element e14.	<i>Escherichia coli</i>
o240#17	625882	625286	199	26.3	-	1012	77.6	196	(T)	<i>ycfA</i>	S18684	PIR Gene T protein	<i>Escherichia coli</i> plasmid p15B

Table 1. Continued.

orf number	start* (nt)	end (nt)	length (aa)	map (min)	direction*	score <sup>c</sup>	% <sup>d</sup>	overlap <sup>e</sup> gene <sup>d</sup>	y gene <sup>f</sup>	accession <sup>g</sup>	product	species
o240#18	626421	625885	179	26.3	-	961	83.1	166	<i>ycfE</i>	S18686	pir Sc/SvMr protein	<i>Escherichia coli</i> plasmid p15B
o241#1	626448	626999	184	26.3	+	1195	100	184 <i>pin</i>		A03545	pir DNA-invertase	<i>Escherichia coli</i>
o241#2	627109	627939	277	26.3	+	1869	100	277 <i>mcrA</i>		A41424	pir Modified cytosine restriction protein A	<i>Escherichia coli</i>
o241#3	628176	628337	54	26.4	+	347	94.5	55 ( <i>icdA, icd, icdE</i> )		P08200	SP Isocitrate dehydrogenase (NADP) (EC 1.1.1.42)	<i>Escherichia coli</i>
o241#4	630822	630094	243	26.4	-	168	32.4	74 ( <i>merR</i> )		P22853	SP Mercuric resistance operon regulatory protein	<i>Bacillus sp.</i> (Strain RC607)
o241#5	632238	631030	403	26.4	-	225	26.1	238		Q04855	SP Hypothetical 80.5 KD protein in <i>nutC</i> 5' region (ORF1).	<i>Azorhizobium caulinodans</i> .
o241#7	634049	635611	521	26.5	+	489	31.6	304 ( <i>ycjC</i> )		P32701	SP Hypothetical 60.8 KD protein in <i>ssb-soxS</i> intergenic region (O528).	<i>Escherichia coli</i>
o242#1	636364	637881	506	26.5	+	289	26.6	458 ( <i>virG</i> )		A32247	pir VirG protein	<i>Shigella flexneri</i>
o242#2	638311	638982	224	26.6	+	865	40.3	372 ( <i>virG</i> )		A32247	pir VirG protein	<i>Shigella flexneri</i>
o242#4	639498	639767	90	26.6	+	165	36.7	90		A43932	pir Mucin	human (fragment)
o242#5	640458	640667	70	26.6	+	370	64.9	77 ( <i>ycjO</i> )		P33924	SP Hypothetical 91.2 KD protein in <i>rplY-narP</i> intergenic region.	<i>Escherichia coli</i>
o242#6	641308	641045	88	26.6	-	555	100	88 <i>minE</i>		P18198	SP Cell division topological specificity factor.	<i>Escherichia coli</i>
o242#7	642124	641315	270	26.6	-	1705	100	270 <i>minD</i>		B31877	pir Cell division inhibitor MinD	<i>Escherichia coli</i>
o242#8	642843	642151	231	26.7	-	1502	100	231 <i>minC</i>		A31877	pir Cell division inhibitor MinC	<i>Escherichia coli</i>
o242#10	644444	644767	108	26.7	+	244	42.7	82 (HI1446)		P44198	SP Hypothetical protein HI1446.	<i>Haemophilus influenzae</i>
o242#11	644842	645498	219	26.7	+	415	38.9	185 ( <i>hpcE</i> )		P37352	SP 2-Hydroxyhepta-2,4-diene-1,7-dioate isomerase (EC 5.3.3.-) (HHDD isomerase) / 5-carboxymethyl-2-oxo-hex-3-ene-1,7-dioate decarboxylase (EC 4.1.1.-) (OpeT decarboxylase).	<i>Escherichia coli</i>
o242#12	645563	646036	158	26.7	+	532	46	150 (HI1355)		P44168	SP Hypothetical protein HI1355.	<i>Haemophilus influenzae</i>
o243#1	647530	647946	139	26.8	+	892	100	139 <i>umuD</i>		A03551	pir UmuD protein	<i>Escherichia coli</i>
o243#2	647949	649214	422	26.8	+	2815	100	422 <i>umuC</i>		B23157	pir UmuC protein	<i>Escherichia coli</i>
o243#3	649799	649266	178	26.8	-	1236	100	178 <i>dsbB</i>		A48288	pir Disulfide bond formation protein DsbB	<i>Escherichia coli</i>
o243#4	651480	649942	513	26.8	-	3196	100	502 <i>nhaB</i>		P27377	SP NA(+)/H(+) antiporte R 2.	<i>Escherichia coli</i>
o244#1	651701	652417	239	26.9	+	1577	99.6	239 <i>fadR</i>		S01288	pir Rregulatory protein FadR	<i>Escherichia coli</i>
o244#2	654004	652475	510	26.9	-	586	31.5	480 ( <i>spoVR</i> )	<i>ycgB</i>	P37875	SP Stage V sporulation protein R.	<i>Bacillus subtilis</i>
o244#3	654334	655629	432	26.9	+	2911	100	432 <i>dadA, dadR</i>		P29011	SP D-amino acid dehydrogenase small subunit (EC 1.4.99.1).	<i>Escherichia coli</i>
o244#4	655642	656709	356	26.9	+	2344	98.9	356 <i>dadX, dadB, ainB</i>		P29012	SP Alanine racemase,catabolic precursor (EC 5.1.1.1).	<i>Escherichia coli</i>
o245#2	659829	660551	241	27.0	+	146	29.8	131 (VII)		P27380	SP Protein P7.	Bacteriophage PRD1
o245#5	664139	662445	565	27.1	-	3884	100	565 <i>treA, osmA</i>		P13482	SP Pperiplasmic trehalase precursor (EC 3.2.1.28)	<i>Escherichia coli</i>
o245#6	665361	664348	338	27.1	-	2267	100	338 <i>IS5</i>		S47725	pir Hypothetical protein F338	<i>Escherichia coli</i>
o245#7	664693	665046	118	27.1	+	778	100	118 <i>IS5</i>		B91483	pir Hypothetical 12K protein (insertion sequence IS5)	<i>Escherichia coli</i>
o245#8	667079	665661	473	27.2	-	330	28.4	334 ( <i>fruB</i> (HI))	<i>ycgC</i>	P23388	SP Multiphosphoryl transfer protein (Mtp) (contains: phosphoenolpyruvate- protein phosphotransferase (EC 2.7.3.9) (Phosphotransferase system, enzyme I); Phosphocarrier protein HPR (protein H); PTS system,	
o245#10	668830	667733	366	27.2	-	510	33	352 ( <i>dhaK</i> )		P45510	SP Dihydroxyacetone kinase (EC 2.7.1.29) (glycerone kinase).	<i>Citrobacter freundii</i>
o245#11	669019	670944	642	27.2	+	3056	71.6	638 ( <i>dhaR</i> )		P45512	SP Glycerol metabolism operon regulatory protein.	<i>Citrobacter freundii</i>
o246#1	673914	671050	955	27.3	-	831	26.6	906 ( <i>prn, omp69A</i> )		P14283	SP Pertactin precursor (outer membrane protein P.69) (P.93).	<i>Bordetella pertussis</i>
o246#2	675771	674686	362	27.4	-	1745	100	268	<i>ychF</i>	P31216	SP Probable GTP-binding protein in <i>pth</i> 3' region (ORF-3) (fragment).	<i>Escherichia coli</i>
o246#3	676475	675894	194	27.4	-	1291	100	194 <i>pth</i>		S16753	pir Aminoacyl-tRNA hydrolase (EC 3.1.1.29)	<i>Escherichia coli</i>
o246#4	676753	677028	92	27.4	+	644	100	92	<i>ychH</i>	P31807	SP Hypothetical 10.5 KD protein in <i>pth-prs</i> intergenic region (ORF-2).	<i>Escherichia coli</i>
o246#6	678738	677089	550	27.4	-	275	25.1	427	<i>ychM</i>	Q02920	SP Early nodulin 70.	<i>Glycine max</i> (Soybean)
o246#7	679834	678893	314	27.5	-	2004	99.7	314 <i>prsA, prs</i>		P08330	SP Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)	<i>Escherichia coli</i>
o246#8	680839	679991	283	27.5	-	1895	100	283	<i>ychB</i>	P24209	SP Hypothetical 30.9 KD protein in <i>hemM-prs</i> intergenic region.	<i>Escherichia coli</i>

Table 1. Continued.

orf number	start* (nt)	end (nt)	length (aa)	map (min)	direction <sup>b</sup>	score <sup>c</sup>	% <sup>d</sup>	overlap <sup>e</sup> gene <sup>f</sup>	y gene <sup>g</sup>	accession <sup>h</sup>	product	species	
o246#9	681462	680842	207	27.5	-	1417	100	207 <i>hemM</i>		A47706	pir	Orf1 3' of <i>hemA</i>	<i>Escherichia coli</i>
o247#1	681676	682929	418	27.5	+	2625	100	418 <i>hemA</i>		P13580	sp	Glutamyl-tRNA reductase (EC 1.2.1.-)	<i>Escherichia coli</i>
o247#2	682974	684053	360	27.5	+	2344	99.7	360 <i>prfA, sueB, uar</i>		P07011	sp	Peptide chain release factor 1 (RF-1).	<i>Escherichia coli</i>
o247#3	684056	684886	277	27.6	+	1368	100	203 <i>hemK</i>		P37186	sp	Possible protoporphyrinogen oxidase (EC 1.3.3.-).	<i>Escherichia coli</i>
o247#5	685333	686088	252	27.6	+	1658	99.2	252	<i>ychA</i>	P20101	sp	Hypothetical 29.0 KD protein in <i>hemK-kdsA</i> intergenic region.	<i>Escherichia coli</i>
o247#6	686127	686978	284	27.6	+	1865	99.6	284 <i>kdsA</i>		A30390	pir	2-dehydro-3-deoxyphosphooxtonate aldolase (EC 4.1.2.16)	<i>Escherichia coli</i>
o247#10	689811	688714	366	27.7	-	2274	100	366 <i>chaA</i>		P31801	sp	Calcium/proton antiporter.	<i>Escherichia coli</i>
o247#11	690081	690308	76	27.7	+	530	100	76 <i>chaB</i>		P39162	sp	Cation transport regulator ChaB.	<i>Escherichia coli</i>
o247#12	690448	691161	238	27.7	+	1608	99.2	238 <i>chaC</i>		P39163	sp	Cation transport protein ChaC.	<i>Escherichia coli</i>
o248#1	691561	691211	117	27.7	-	739	99.1	117	<i>ychN</i>	P39164	sp	Hypothetical 12.7 KD protein in <i>chaC-narL</i> intergenic region.	<i>Escherichia coli</i>
o248#2	691887	693137	417	27.7	+	698	30	480	<i>ychP</i>	P11922	sp	Invasin.	<i>Yersinia pseudotuberculosis</i>
o248#3	693791	693144	216	27.8	-	1357	100	216 <i>narL, frdR</i>		P10957	sp	Nitrate/nitrite response regulator protein <i>narL</i> .	<i>Escherichia coli</i>
o248#4	695580	693787	598	27.8	-	3892	100	598 <i>narX, narR</i>		P10956	sp	Nitrate/nitrite sensor protein NarX (EC 2.7.3.-).	<i>Escherichia coli</i>
o248#5	695919	697307	463	27.8	+	3055	100	463 <i>narK</i>		S05239	pir	Nitrate transport protein NarK	<i>Escherichia coli</i>
o249#1	697829	701566	1246	27.9	+	8187	96.8	1246 <i>narG, narC, bisD</i>		P09152	sp	Respiratory nitrate reductase 1 alpha chain (EC 1.7.99.4).	<i>Escherichia coli</i>
o249#3	701566	703101	512	27.9	+	3414	96.1	512 <i>narH</i>		JV0038	pir	Nitrate reductase (EC 1.7.99.4) beta chain	<i>Escherichia coli</i>
o249#4	703101	703808	236	28.0	+	1533	100	236 <i>narJ</i>		B27737	pir	NarJ protein	<i>Escherichia coli</i>
o249#5	703811	704485	225	28.0	+	1522	100	225 <i>narL, chll</i>		P11350	sp	Respiratory nitrate reductase 1 gamma chain (EC 1.7.99.4) (cytochrome b-nr).	<i>Escherichia coli</i>
o249#6	705150	705052	33	28.0	-	232	100	33 <i>tpx</i>		A90813	pir	Protamine-like protein	<i>Escherichia coli</i>
o249#7	706586	705747	280	28.0	-	1844	100	280 <i>purU, tgs</i>		P37051	sp	Formyltetrahydrofolate deformylase (EC 3.5.1.10)	<i>Escherichia coli</i>
o249#8	707094	706639	152	28.1	-	156	48.1	54 ( <i>secA</i> )	<i>ychJ</i>	S47149	pir	SecA protein	<i>Staphylococcus carnosus</i>
o249#9	707168	708109	314	28.1	+	2061	99.7	314	<i>ychK</i>	P37053	sp	Hypothetical 34.4 KD protein in <i>hnr-purU</i> intergenic region.	<i>Escherichia coli</i>
o249#10	708204	709214	337	28.1	+	2213	100	337 <i>hnr</i>		A36871	pir	Regulator response protein homolog.37.3K	<i>Escherichia coli</i>
o249#11	709419	710324	302	28.1	+	1976	100	302 <i>galU</i>		A55585	pir	UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9)	<i>Escherichia coli</i>
o250#1	710884	710474	137	28.1	-	853	100	137 <i>osmZ, bglY</i>		S28633	pir	DNA-binding protein H-NS	<i>Escherichia coli</i>
o250#2	711489	712103	205	28.2	+	1358	100	205 <i>tdk</i>		S15206	pir	Thymidine kinase (EC 2.7.1.21)	<i>Escherichia coli</i>
o250#3	712978	712391	196	28.2	-	267	38.4	125 (IS4)	<i>ychG</i>	P03835	sp	Insertion element IS4 hypothetical 50.4 KD protein.	<i>Escherichia coli</i>
o250#4	716083	713411	891	28.2	-	5736	100	891 <i>adhE</i>		JS0406	pir	Alcohol dehydrogenase (EC 1.1.1.1)	<i>Escherichia coli</i>
o250#5	716659	717204	182	28.3	+	1161	100	182	<i>ychE</i>	P25743	sp	Hypothetical 19.7 KD protein in <i>adhE-oppA</i> intergenic region.	<i>Escherichia coli</i>
o250#6	717515	717877	121	28.3	+	773	100	121		JQ0039	pir	Hypothetical 13K protein (insertion sequence IS2)	<i>Escherichia coli</i>
o250#7	717838	718122	95	28.3	+	637	98.9	95 IS2		P19777	sp	Insertion element IS2 Hypothetical 34.4 KD protein (ORF2).	<i>Escherichia coli</i>
o250#8	718122	717886	79	28.3	-	525	100	79 IS2		P19779	sp	Insertion element IS2 Hypothetical 16.4 KD protein (ORF4).	<i>Escherichia coli</i>
o250#9	717911	718122	70	28.3	+	480	98.6	70 IS2		P19778	sp	Insertion element IS2 Hypothetical 11.1 KD protein (ORF3).	<i>Escherichia coli</i>

a. The nucleotide residues are numbered from the left end of Kohara mini set clone #162

b. +, clockwise; -, counterclockwise direction of transcription in the *E. coli* genetic map

c. FASTA optimal score in the overlapped region

d. amino acid identity in the overlapped region

e. size of the region where significant similarity is detected

f. Gene names in parentheses indicate that ORFs identified are similar in amino acid sequence to those gene products

g. hypothetical genes designated by K.E. Rudd et al.<sup>4</sup>

h. sp, SWISS-PROT; pir, NBRF-PIR protein database

