Quick identification of Type I restriction enzyme isoschizomers using newly developed pTypeI and reference plasmids

Junichi Ryu[1] and Edward Rowsell[2],*

[1]Division of Microbiology and Molecular Genetics, Department of Biochemistry and Microbiology and [2]Department of Pathology and Laboratory Medicine, School of Medicine, Loma Linda University, Loma Linda, CA92350, USA

ABSTRACT

Although DNA-recognition sequences are among the most important characteristics of restriction enzymes and their corresponding methylases, determination of the recognition sequence of a Type-I restriction enzyme is a complicated procedure. To facilitate this process we have previously developed plasmid R-M tests and the computer program RM search. To specifically identify Type-I isoschizomers, we engineered a pUC19 derivative plasmid, pTypeI, which contains all of the 27 Type-I recognition sequences in a 248-bp DNA fragment. Furthermore, a series of 27 plasmids (designated 'reference plasmids'), each containing a unique Type-I recognition sequence, were also constructed using pMECA, a derivative of pUC vectors. In this study, we tried those vectors on 108 clinical *E. coli* strains and found that 48 strains produced isoschizomers of Type I enzymes. A detailed study of 26 strains using these 'reference plasmids' revealed that they produce seven different isoschizomers of the prototypes: EcoAI, EcoBI, EcoKI, Eco377I, Eco646I, Eco777I and Eco826I. One strain EC1344 produces two Type I enzymes (EcoKI and Eco377I).

INTRODUCTION

Restriction enzymes and their corresponding methylases (R-M enzymes) are produced by various bacterial species and are traditionally classified into three different types (1). Recently, Roberts *et al.* (2) have proposed a new class IV for restriction enzymes specific to methylated DNA. Although close to 4000 restriction enzymes are already known, current genome projects reveal the possible existence of many additional R-M systems in diverse bacterial species (3).

Identification of the recognition sequences is relatively simple for Type II enzymes, since either the crude extracts or partially purified Type II enzymes cut DNA at distinctive locations, resulting in fixed lengths of fragments. Subsequently, computer programs help to determine the possible cutting sites from the fragments (3,4). This convenient method, however, cannot be used for Type I enzymes. These enzymes do not produce any distinctive DNA fragments in the crude extracts because those enzymes cut DNA far from the recognition sequences and at random positions. Thus, no simple method to identify recognition sequences has yet been developed for Type I or Type III enzymes. One alternative approach is to identify the binding site of the companion methylase, which is shared by the restriction endonuclease (5). However, characterizing the methylation site requires extensive enzyme purification, isotope labeling and the application of a computer program to analyze the results. This protocol is not suitable for quick screening of many bacterial strains.

As shown by the extensive list of the isoschizomers of the Type II enzymes (273 prototypes and more than 3400 isoschizomers) (3), finding new prototype enzymes has rapidly become more difficult. It is possible that many of the more than 1000 putative Type II enzymes found through genome analysis are also isoschizomers. On the other hand, only a handful of active Type I and Type III enzymes (60 and 7, respectively) have been discovered so far. Among the 60 Type I enzymes, the recognition sequences of only half (32 enzymes) have been identified. This includes 27 naturally occurring recognition sequences and five artificially constructed recognition sequences. From various bacterial genome projects about 1200 putative Type I enzymes and about 300 putative Type III

enzymes have been identified (3). Thus, there are many more Type I and Type III restriction enzymes to be characterized.

A simple and economical method to identify Type I and Type III recognition sequences is needed. As described elsewhere (6), we have applied the classical phage EOP (Efficiency of Plating) assay to a chemical plasmid transformation protocol and developed plasmid R-M tests to detect the presence of restriction enzymes in bacterial cells. With these tests, together with the introduction of a new computer program, RM search (7), we have developed a novel method to identify their recognition sequences (6). This method was used to identify several recognition sequences for Type I enzymes in bacteria belonging to the *Klebsiella* and *Salmonella* genera (8,9). It was also used to find new restriction enzymes in clinical *E. coli* strains (6). The application of this method resulted in a doubling of the number of known recognition sequences for the naturally occurring Type I enzymes from 14 to 27 and two EcoBI isoschizomers were found (3). Our data also showed that this plasmid transformation method is useful for detecting not only the activities of Type I enzymes but also those of Type II and Type III enzymes (6). Nevertheless, because of the complexity of the recognition sequences of the Type III enzymes, the elucidation of their recognition sequences using the RM search program has been so far limited to Type I and Type II enzymes (7).

Although the total number of known Type I recognition sequences is only 32 and the potential to find many more new recognition sequences is high at present, we still expect to encounter many isoschizomers in the process of screening for Type I enzymes. To quickly distinguish new prototypes from isoschizomers, we have developed here a modified pUC19 plasmid, pTypeI, which contains all the presently known Type I recognition sequences. If a test bacterial strain produces one of the known Type I enzymes, they will restrict pTypeI on transformation and as a result, produce fewer transformants than a control plasmid without the recognition sequence. Each of the isoschizomer sequences can be estimated using a plasmid chart developed in this project and then can be confirmed by using the complete set of reference plasmids developed here. We have performed a feasibility study on these plasmids using *E. coli* strains readily available from our hospital and show here the usefulness of the plasmids by identifying 27 Type I isoschizomers.

MATERIALS AND METHODS

Bacterial strains

Bacteria producing known Type I restriction enzymes (Table 1) were all from our lab collection. Approximately 700 clinical *E. coli* strains were obtained originally from the clinical laboratory of Loma Linda University Medical Center between 2000 and 2004. Some of them were previously used to identify new Type I enzymes (6,10). In this study, 108 ampicillin-sensitive *E. coli* strains, a subset of those clinical samples, were used for the screening of Type I isoschizomers. Some of the *E. coli* strains

Table 1 Plasmid restriction tests using pTypeI for known R-M systems

Strains	Type I system	Restriction EOT (±STD)	Source or reference
E. coli system			
2739	EcoAI	4.4×10^{-2} ($\pm 3 \times 10^{-2}$)	Lab stock
L4001	EcoBI	4.8×10^{-2} ($\pm 7 \times 10^{-2}$)	Lab stock
1225	EcoKI	3.3×10^{-2} ($\pm 5 \times 10^{-2}$)	Lab stock
EC377	Eco377I	1.1×10^{-1} ($\pm 2 \times 10^{-1}$)	(6)
EC585	Eco585I	1.1×10^{-1} ($\pm 1 \times 10^{-1}$)	(6)
EC646	Eco646I	4.9×10^{-2} ($\pm 3 \times 10^{-2}$)	(6)
EC777	Eco777I	6.7×10^{-3} ($\pm 1 \times 10^{-2}$)	(6)
EC394	Eco394I	2.7×10^{-3} ($\pm 3 \times 10^{-3}$)	(10)
EC912	Eco912I	1.5×10^{-2} ($\pm 2 \times 10^{-2}$)	(10)
Salmonella system			
L4029	StySEAI	7.2×10^{-2} ($\pm 6 \times 10^{-2}$)	(11)
L4039	StySGI	9.4×10^{-2} ($\pm 1 \times 10^{-1}$)	(11)

EOT (Efficiency of Transformation) is number of transformants with pTypeI/number of transformants with pMECA. Except for EcoKI, the pMECA plasmid does not contain the relevant recognition sequences for the test strains. In case of EcoKI, an EcoKI modified pMECA was used as a control. Each value is the average of three experiments.

containing *Salmonella* Type I restriction enzymes (*E. coli–Salmonella* hybrid) (11) were also used in this study.

pTypeI plasmid

The 248-bp insertion fragments of the pTypeI plasmid was synthesized by Midland Certified Reagent Company (Midlands, Texas) and inserted at the HincII site of pUC19 (12) (Figure 1). This DNA fragment contains 27 naturally occurring Type I recognition sequences overlapping each other. It consists of two sub-fragments; fragments A (117 bp) which contains the 17 recognition sequences found so far in *E. coli* and fragment B (113 bp) which contains the remaining 10 Type I recognition sequences found in other bacteria. When the prototype recognition sequence contains degenerate bases, only one base was selected, thus only one sequence is represented for each prototype. In addition to the recognition sequences, fragment B contains a portion of the MCS (multiple cloning site) of pUC19 in such a way that the entire section, which consists of a 'Type I sequence cassette' (about 250 bp, Figure 1), can be excised from the pTypeI plasmid and transferred to any other vectors using restriction enzymes such as EcoRI, KpnI and BamHI. Thus, once subcloned into plasmids able to replicate in other genera, this fragment can be used to characterize Type I systems in a variety of non-enteric bacteria. The EcoRV sites are located at both ends of the 248-bp fragment. In addition, fragments A and B are flanked by a unique EcoRV site. This allows ligation of each fragment to any blunt end cloning sites.

Reference plasmids

To confirm the predicted pattern for each Type I recognition sequence, a series of plasmids each containing a single Type I recognition site have been constructed. These plasmids were termed collectively as the reference plasmids (9). An example of the reference plasmid with the
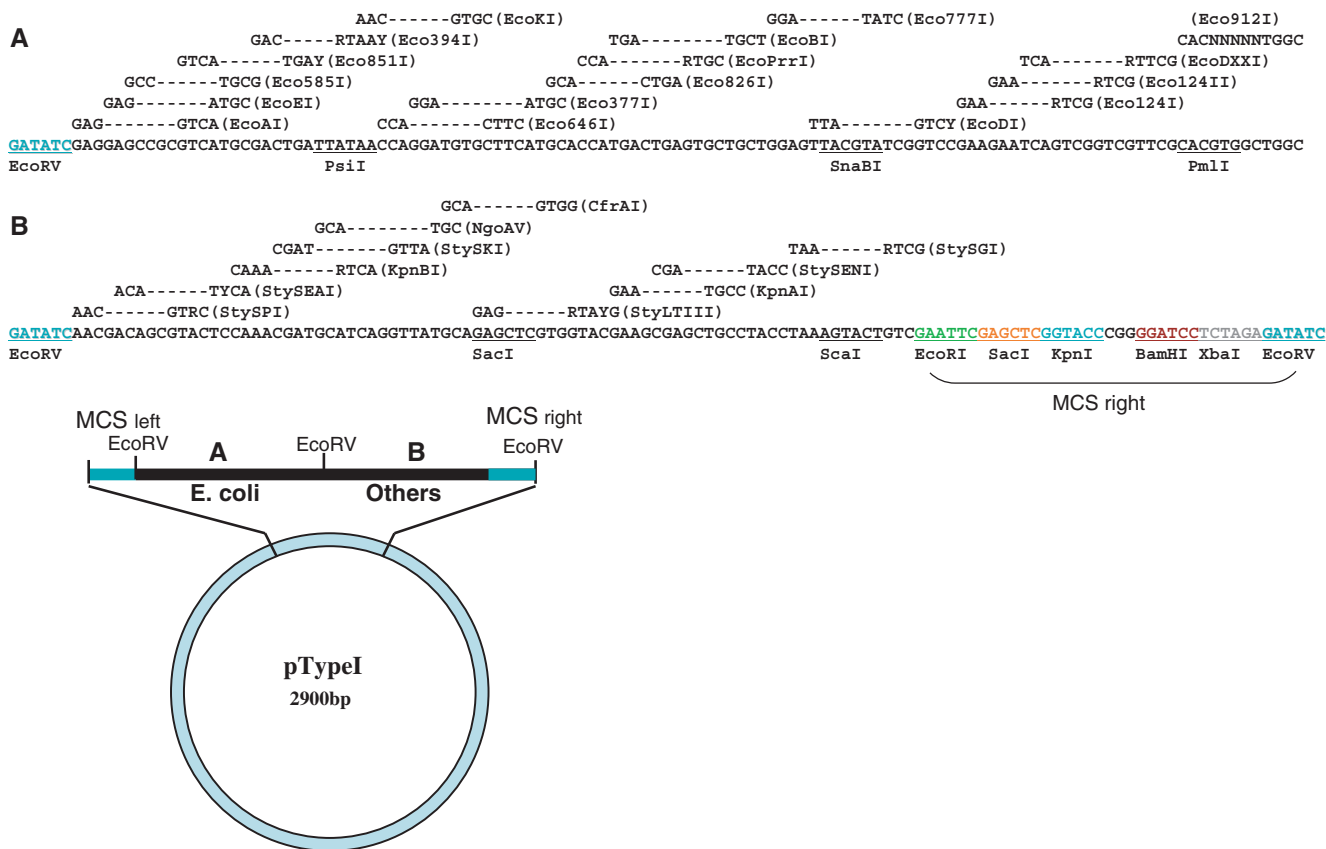
Figure 1. pTypeI plasmid and relevant DNA sequence. A 248-bp fragment containing all 27 naturally occurring Type I sequences was cloned in the HincII site (blunt end) of pUC19. This insert contains 17 Type I sequences (Fragment **A**) found in *E. coli* and 10 Type I sequences (Fragment **B**) found in other bacteria. The 248-bp insert starts and ends at an EcoRV site. Another EcoRV site is located at the boundary of fragments A and B such that those two fragments can be easily subcloned into other vectors. The left side (MCS left) contains half of the original MCS sequences of pUC19: EcoRI, SacI, KpnI, SmaI, XmaI, BamHI, XbaI and ends at the remnant of HincII (GTC)–the cloning site. This MCS left portion is shown above as a linear bar extending from the 248-bp insertion fragment, rather than as a part of the circular plasmid. A nearly identical MCS fragment was added to the right end of the synthetic fragment (MCS right), where restriction sites are colored. Therefore, the entire group of Type I sequences can be cut out as a cassette using any of those enzymes except for SacI. The right end of the 248-bp fragment is an EcoRV site which is connected to the cleaved other half of the HincII site (GAC), thus the original HincII site was lost in this cloning process. In the design of the 248-bp fragment, a few convenient unique restriction sites were added. These are PsiI, SnaBI, PmlI, and ScaI (all underlined).

EcoBI sequence, pEcoBI, is shown in Figure 2. Since we reported previously about a half (13/27) of those reference plasmids for Type I enzymes (6–10), only the remaining reference plasmids were constructed here. These are pCfrAI, pEcoAI, pEcoDI, pEcoXXI, pEcoEI, pEcoKI, pEco124I, pEco124II, pEcoprrI, pNgoAV, pStyLTIII, pSKI and pSPI, each containing the recognition sequence for the Type I enzymes indicated in its name.

Oligonucleotide production and DNA sequencing were carried out at the core facility in the Center for Molecular Biology, Loma Linda University.

Plasmid R-M tests

Plasmid restriction and modification (R-M) tests and the mini-scale (10 μl) cold CaCl$_2$-heat shock transformation method were described previously in detail (6). In short, a plasmid with an unmethylated recognition site will be restricted by a restriction enzyme in the transformation process, whereas a plasmid with no recognition site will transform normally and produce the full number of transformants on a selection plate such as LB agar with

ampicillin (Restriction test). A surviving plasmid from the restriction test is recovered. This plasmid is fully modified (methylated) and suffers no more restriction in the following transformation of the restricting bacteria (Modification test). Plasmids used here (Table 2) are various subclones of phage lambda (pL series) or portions of the *E. coli* chromosome (pE series) cloned in either pUC19 (12) vector or its derivative, pMECA (10), as described previously in detail (6,9). Unmethylated plasmids were prepared from the strain GM272 (Dam$^-$, Dcm$^-$, R^-_KM^-_K). About 10 ng of DNA was used for each transformation and ampicillin resistant colonies were selected on L-amp agar plates.

RESULTS

Restriction Modification (R-M) tests of pTypeI plasmid using known prototype strains

The plasmid R-M tests were applied to the newly constructed pTypeI plasmid and 11 bacterial strains (nine *E. coli* and two *Salmonella–E.coli* hybrid strains),
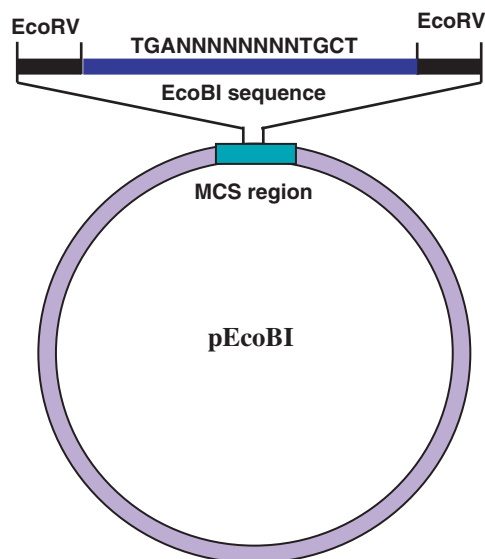
Figure 2. A diagram illustrating the design of a reference plasmid, pEcoBI, as an example of the reference plasmid construction. In the actual experiment, a 21mer (ATCTGAGTACGCGTTGCTGAT) that contains a single EcoBI sequence (underlined) was cloned in the single EcoRV site (blunt end) located in the MCS site in pMECA (13). The final product is an EcoBI sequence flanked by EcoRV sequences. This sequence also contains a MluI sequence (ACGCGT) in the random sequence (NNNNNNNN) region for easy identification of the cloned fragment.

available in our laboratory, each producing unique Type I prototype enzymes. Either pUC19 or its derivative pMECA (10) were used as negative controls. The relative efficiency of transformation (EOT) values are unique to each strain and dropped from 1.0 to 10^{-1} to 10^{-3} level when the pTypeI plasmid was used (Table 1). This level of restriction was observed previously with other plasmids containing a single target recognition site (6,9). To confirm that this reduction is due to the classical restriction-modification phenomenon, plasmid modification tests (6) were also performed using the plasmids obtained from surviving Amp^R transformants. As expected, modified plasmids were all resistant to restriction and showed EOT values close to 1.0 (0.9–1.7). These results show the usefulness of pTypeI in identifying known Type I systems in various bacterial strains.

Screening of isoschizomers in clinical *E. coli* strains

To test the effectiveness of pTypeI with wild-type bacteria, we chose *E. coli* clinical strains obtained from Loma Linda University Medical Center. Using the quick CaCl$_2$ heat-shock transformation method (6), 108 ampicillin sensitive strains were tested for their transformation ability with the pTypeI plasmid. The plasmid pMECA was used as a negative control. Frequencies of transformation varied from strain to strain and as a result, a wide range of Amp^R colonies (from a few to 10^4) per plate were observed. Among the 108 strains tested, 48 (44%) showed less than 20% (EOT < 0.2) of the transformant numbers on pTypeI compared to the control pMECA. This suggests that

about half of the 108 clinical strains are expressing known Type I restriction enzyme activities.

Identification of isoschizomers

To identify the isoschizomers, we selected 26 strains among 48 possible Type I enzyme producers that fulfill the following two criteria; (i) they produced at least 10 transformant colonies on control plates (pMECA) in our standard plasmid restriction test; (ii) they showed clear-cut reduction of transformant numbers (EOT < 0.1) when transformed with pTypeI.

To test those 26 candidate strains, we chose 13 plasmids from among the plasmids we used previously (6,9,10) to distinguish one isoschizomer producer from another. The size of the plasmids ranged from 3 kb (pL23 and pL37) to 20 kb (pL2), but most were between 5 and 9 kb. Table 2 shows the presence (+) or absence (−) of the individual recognition sequence on the test plasmids. Those test plasmids were chosen so that each prototype will restrict a unique combination of test plasmids showing a unique +/− pattern. Comparing this pattern with the plasmid restriction test of the clinical *E. coli* strains gives us a prediction for each isoschizomer. The 26 *E. coli* clinical strains described above were then subjected to the plasmid restriction test. The +/− pattern of each clinical strain was matched to one of the patterns in Table 2. Thus the identities of seven different isoschizomers were predicted from 25 clinical strains (Table 2, last column). However, the pattern of the strain EC1344 (labeled '*') did not fit any of the patterns in Table 2, suggesting the presence of two or more restriction enzymes (see below).

To confirm the predicted recognition sequence (Table 2), each clinical strain was tested with the corresponding reference plasmid (Figure 2). Plasmid modification tests were also performed for further confirmation. A typical example of this confirmation test obtained for strain EC1446 (an EcoAI producer) is shown in Table 3. In this example, relative efficiency of transformation (EOT) values were compared to that of the control pMECA. Non-modified (non-methylated) plasmids were severely restricted (center two columns—restriction test), whereas modified (methylated) plasmids obtained from EC1446 strain were not restricted (right two columns—modification test). Similar experiments were performed for each strain and the predictions from the Table 2 were all confirmed (data not shown).

As described above, the strain EC1344 did not fit completely to any of the patterns in the Table 2. This strain was tested using each of the 27 reference plasmids. Interestingly, reductions of EOT values were observed in two reference plasmids, pEco377I and pEcoKI. Modification tests confirmed the presence of two R-M systems (Eco377I and EcoKI) in strain EC1344, thus Table 2 has a double entry for strain EC1344. Since *E. coli* strains generally contain only one Type I enzyme in the fixed genetic location at 98.5 min (so called immigration control region), this unusual phenotype bears further investigation.

Table 2 Type I recognition sequences and the presence (+) or absence (−) of the recognition site in the test plasmid set. Assignment of the clinical strains (EC) possessing each system

Type I enzymes	Recognition sequences	Plasmids in the Test Set													EC strains tested positive
		pL1	pL2	pL3	pL4	pL5	pL6	pL8	pL10	pL13	pL19	pL23	pL37	pE19	
E. coli system															
EcoAI	GAGNNNNNNNGTCA	−	−	−	−	+	−	−	−	−	+	−	−	−	1446
EcoBI	TGANNNNNNNNTGCT	+	+	−	−	−	+	+	−	−	−	−	−	−	1441, 1443, 1444, 1447
EcoDI	TTANNNNNNNGTCY	−	−	+	+	+	+	−	+	−	+	−	−	−	
EcoDXXI	TCANNNNNNNRTTC	−	+	−	+	+	+	+	+	+	+	−	+	+	
EcoEI	GAGNNNNNNNATGC	+	+	−	+	+	+	+	+	−	−	−	−	+	
EcoKI	AACNNNNNNGTGC	−	+	−	−	+	+	+	−	−	−	−	−	+	1342, 1344*, 1394,
EcoR124I	GAANNNNNNRTCG	+	+	+	+	+	+	+	+	+	+	−	−	+	
EcoR124II	GAANNNNNNNRTCG	−	+	+	+	+	+	+	+	−	+	−	−	+	
EcoprrI	CCANNNNNNNRTGC	−	+	−	−	+	−	−	−	−	+	−	+	−	
Eco377I	GGANNNNNNNNATGC	+	+	+	+	+	+	+	−	−	−	−	−	−	1323, 1344*, 1386, 1427, 1430, 1456, 1476
Eco394I	GACNNNNNRTAAY	+	+	+	−	+	−	−	−	−	−	−	−	−	
Eco585I	GCCNNNNNNNTGCG	−	+	−	−	+	−	+	−	−	−	−	−	+	
Eco646I	CCANNNNNNNCTTC	−	+	−	−	+	+	+	−	−	+	−	−	+	1341, 1393, 1413, 1422, 1424, 1432,
Eco777I	GGANNNNNNTATC	−	+	+	+	+	+	+	+	+	+	+	−	+	1345, 1348, 1360, 1412, 1467
Eco826I	GCANNNNNNCTGA	+	+	−	−	+	+	+	−	−	+	−	−	+	1455
Eco851I	GTCANNNNNNNTGAY	+	+	−	−	−	+	+	−	−	−	−	−	+	
Eco912I	CACNNNNNTGGC	+	+	−	+	−	−	+	+	−	−	−	−	+	
Other bacterial system															
StyLTIII	GAGNNNNNNNRTAYG	+	+	−	+	−	+	+	+	−	−	−	−	+	
StySEAI	AGANNNNNNYTTRG	−	−	−	−	+	+	−	−	−	−	−	−	−	
StySENI	CGANNNNNNTACC	−	+	−	+	−	−	−	+	−	−	−	−	+	
StySGI	TAACNNNNNNNNRTCG	−	+	−	−	+	−	−	−	−	+	−	−	+	
StySKI	CGATNNNNNNNGTTA	−	+	−	−	+	−	−	−	−	+	−	−	−	
StySPI	AACNNNNNNNGTRC	+	+	−	+	+	+	+	−	−	−	−	+	+	
KpnAI	GAANNNNNNNTGCC	+	+	−	+	−	+	+	−	−	+	+	−	−	
KpnBI	CAAANNNNNNRTCA	−	+	+	+	−	+	−	+	−	−	−	+	+	
CfrAI	GCANNNNNNNNGTGG	+	+	+	+	+	−	+	−	+	−	−	−	+	
NgoAV	GCANNNNNNNNTGC	+	+	+	+	+	+	+	+	−	−	+	−	+	

Table 3 Plasmid R-M tests[a] of EC1446 strain

Bacteria	pMECA	Restriction test		Modification test	
		pTypeI.0	pEcoAI.0	pTypeI.1446	pEcoAI.1446
E. coli C	1.0	1.0	1.0	1.0	1.0
EC1446	1.0	7.0×10^{-3}	1.4×10^{-2}	0.9	1.3

Relative EOT values are shown.
Following the tradition used for modified and un-modified phages for EOP assay, a symbol (.0) was added to the un-modified plasmids, whereas for modified (methylated) plasmids prepared from the EC1446 strain, a symbol (.1446) was added. Amp[R] transformant numbers were normalized first to *E. coli* C (non restriction enzyme producer) and then to pMECA. Actual transformant numbers observed were between 1000 to 3000 for *E. coli* C and 100 to 300 for EC1446.
[a]For the Restriction test, plasmids were harvested from strain GM272 (*dam*⁻, *dcm*⁻ *hsdS*$_K$⁻) and pMECA was used as a control.

In summary, a total of 27 isoschizomers were identified among 26 *E. coli* strains (Table 2). Without exception, all 26 strains produce one or two Type I enzymes. The two most common isoschizomers are Eco377I and Eco646I, which between them account for 48% (13/27) of the R-M systems identified in these strains. The classical Type I enzymes (EcoBI, EcoKI and EcoAI) accounted for only 30% (8/27) in these clinical samples.

DISCUSSION

After identifying the restriction activities in bacterial cells, it is important to know whether the restriction enzyme in question recognizes a new prototype sequence or the same sequence as already reported (isoschizomer). We have developed here a new convenient tool to facilitate this process for Type I restriction enzymes and readily identified several isoschizomers. Isoschizomers are important for evolutionary and population studies. For the purpose of screening for new prototype enzymes, however, they should be discarded swiftly.

In this survey, the simple plasmid restriction tests revealed that nearly half of newly screened clinical *E. coli* had known Type I enzymes. The remaining strains await further investigation; they may produce either no restriction enzyme, or produce new Type I or other Types (II or III) of restriction enzymes. A previous survey of Type II restriction enzymes in *E. coli* revealed Type II isoschizomers at least in 25% (219/864) of the strains examined (14).

The engineering of pUC19 to add unique recognition sites for all the Type I sequences was possible simply because each of the Type I recognition sequences reported contain unique seven or eight nucleotides (six in the case

of NgoAV) which, so far, occur sporadically in pUC 19. Among the 27 Type I prototypes, however, there are four cases (EcoKI, EcoprrI, StySPI, StySENI) where the recognition sequences have occurred once, and one case (StyLTIII) where it occurred three times in pUC19 and pMECA. In the former cases, the isoschizomers (for example, EcoKI) could be identified by comparing the EOT values of the pUC19 and pTypeI, since a much reduced EOT can be expected with pTypeI (two EcoKI sites) compared to pUC19 (one EcoKI site). However, distinguishing a plasmid with three recognition sites from four sites is difficult (6), we would have probably missed StyLTIII isoschizomers in the present study. These Type I sites in the vectors could be eliminated by mutations, or the fragment can be cloned into another vector lacking the relevant sites. Alternatively, methylated pUC19 can be prepared in the proper host and used as a negative control.

A new specificity would be also missed if it recognized a sequence in the plasmid backbone. This would reduce overall transformation efficiency equally for test and control plasmid. This will affect a minority of possible Type I specificities.

The Type I systems discovered in this study (total 27) represent a small fraction of local clinical *E. coli* samples. Nevertheless, we observed 7 out of the 17 known *E. coli* Type I systems in these samples, including classical EcoKI, EcoBI and EcoAI systems. Screenings for Type I enzymes in *E. coli* and other enteric bacteria were previously performed by DNA–DNA hybridization using probes from cloned DNA for the enzymes, EcoKI, EcoAI (15–17) and StySBLI (16,17). Since DNA hybridization studies do not necessarily provide a DNA-recognition sequence, it is beneficial to use our present method to obtain this critical information. Thus, it would be extremely interesting to revisit the ECOR collection (17) to search for Type I restriction enzymes, since it represents *E. coli* samples with a wide geographical scale and many different origins.

The strain EC1344 seems to contain two independent R-M systems, EcoKI and Eco377I. The mapping of the Eco377I locus relative to the established EcoKI locus (98.5 min region) in this strain and other strains producing only the Eco377I enzyme are crucial for the further characterization of the strains producing Eco377I. It is also interesting to examine the prevalence of Eco377I in clinical samples and any possible toxin production or connection to diseases. Recently, the Eco377I system was also found in an *E. coli* strain, AO34/86, which is efficiently used in the prophylaxis and treatment of nosocomial infections and diarrhea of preterm and new-born infants in Czech pediatric clinics (M. Weiserova, personal communication). The Eco377I system belongs to the same Type IB family (represented by EcoAI) as the notoriously toxic *E. coli* strain, EcoAO157:H7.

It is very likely that in the near future, more Type I prototype sequences will be identified. In such a case, another plasmid (pTypeI#2) will be designed and used similarly as pTypeI. In fact, a 'super pTypeI' plasmid could be constructed in the future to detect an exhaustive list of the set of Type I prototypes. If, as with Type II enzymes, there were about 270 Type I prototypes, the required DNA insert would still be less than 2-kb long.

Type I enzymes seem to have spread throughout the bacterial kingdom and Archaea (3). This plasmid R-M test, however, has an obvious limitation in that it can only be used in the strains transformable by chemical methods or electroporation. To expand this plasmid R-M test for Type I enzyme screening from enteric bacteria to other gram negative or to gram positive bacteria, the 248-bp Type I cassette will need to be subcloned either into broad host range vectors, or shuttle vectors able to replicate both in *E. coli* and each bacterial group being tested. An alternative would be to clone the R-M genes from a bacterium of interest into an *E. coli* vector and examine the recognition sequences in *E. coli*, since numerous *E. coli* plasmids have already been developed for this purpose (6,8,9). In such a case, it is possible that the expression of the R-M genes would need to be optimized with respect to the codon usage and other factors. Since the DNA and protein sequences are already known, such an approach would be particularly useful for the R-M genes already found in various bacterial genomes.

REFERENCES

1. Bickle,T.A. and Kruger,D.H. (1993) Biology of DNA restriction. *Microbiol.Rev.*, **57**, 434–450.
2. Roberts,R.J., Belfort,M., Bestor,T., Bhagwat,A.S., Bickle,T.A., Bitinaite,J., Blumenthal,R.M., Degtyarev,S.K., Dryden,D.T., Dybvig,K. *et al.* (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.*, **31**, 1805–1812.
3. Roberts,R.J., Vincze,T., Posfai,J. and Macelis,D. (2007) REBASE - enzymes and genes for DNA restriction and modification. *Nucleic Acids Res.*, **35**, D269–D270.
4. Gingeras,T.R., MIlazzo,J.P. and Roberts,R.J. (1978) A computer assisted method for the determination of restriction enzyme recognition sites. *Nucleic Acids Res.*, **5**, 4105–4127.
5. Nagaraja,V., Stieger,M., Nager,C., Hadi,S.M. and Bickle,T.A. (1985) The nucleotide sequence recognized by the *Escherichia coli* D Type I restriction and modification enzyme. *Nucleic Acids Res.*, **13**, 389–399.
6. Kasarjian,J.K.A., Iida,M. and Ryu,J. (2003) New restriction enzymes discovered from *Escherichia coli* clinical strains using a plasmid transformation method. *Nucleic Acids Res.*, **31**, e22.
7. Ellrott,K.P., Kasarjian,J.K.A., Jiang,T. and Ryu,J. (2002) Restriction enzymes recognition sequence search program. *Biotechniques*, **33**, 1322–1326.
8. Chin,V., Valinluck,V., Magaki,S. and Ryu,J. (2004) KpnBI is the prototype of a new family (IE) of bacterial Type I restriction-modification system. *Nucleic Acids Res.*, **32**, e138.
9. Kasarjian,J.K.A., Hidaka,M., Horiuchi,T., Iida,M. and Ryu,J. (2004) The recognition and modification sites for the bacterial Type I restriction systems KpnAI, StySEAI, StySENI and StySGI. *Nucleic Acid Res.*, **32**, e82.

10. Kasarjian,J.K.A., Kodama,Y., Iida,M., Matsuda,K. and Ryu,J. (2005) Four new Type I restriction enzymes identified in *Escherichia coli* clinical isolates. *Nucleic Acid Res.*, **33**, e114.

11. Bullas,L.R., Colson,C. and Neufeld,B. (1980) Deoxyribonucleic acid restriction modification systems in *Salmonella*: chromosomally located systems of different serotypes. *J. Bacteriol.*, **141**, 275–292.

12. Yanisch-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.

13. Thompson,J.M. and Parrott,W.A. (1998) pMECA: A Cloning Plasmid with 44 Unique Restriction Sites that Allows Selection of Recombinants Based on Colony Size. *Biotechniques*, **24**, 922–928.

14. Janulaitis,A., Kazlauskiene,R., Lazareviciunte,L., Gilvonauskaite,R., Steponaviciene,D., Jagelavicius,M., Petrusyte, Bitinaite,J., Vizeviciute,Z., Kiuduliene,E. *et al*. (1988) Taxonomic specificity of restriction-modification enzymes. *Gene*, **74**, 229–232.

15. Daniel,A., Fuller-Pace,F., Legge,D. and Murray,N. (1988) Distribution and Diversity of *hsd* genes in *Escherichia coli* and Other Enteric Bacteria. *J. Bacteriol.*, **170**, 1775–1782.

16. Barcus,V., Titheradge,A. and Murray,N. (1995) The Diversity of Alleles at the *hsd* Locus in Natural Population of *Escherichia coli*. *Genetics*, **140**, 1187–1197.

17. Ochman,H. and Selander,R.K. (1984) Standard Reference Strains of *Escherichia coli* from Natural Populations. *J. Bacteriol.*, **157**, 690–693.