Epidemiol. Infect. (2003), **132**, 61–65. © 2004 Cambridge University Press
DOI: 10.1017/S095026880300133X Printed in the United Kingdom

A combined AFLP–multiplex PCR assay for molecular typing of *Escherichia coli* strains using variable bacterial interspersed mosaic elements

R. HORVÁTH[1], M. DENDIS[1], J. SCHLEGELOVÁ[2], F. RŮŽIČKA[3] AND J. BENEDÍK[4]*

[1] *Genex CZ Ltd, Podstránská 74, 600 00 Brno, Czech Republic*
[2] *Veterinary Research Institute, Hudcova 70, 621 00 Brno, Czech Republic*
[3] *Institute of Microbiology, Medical School, Masaryk University and St Anna University Hospital Brno, Pekařská 53, 635 00 Brno, Czech Republic*
[4] *Genotyp Ltd, Babičkova 116/29, 613 00 Brno, Czech Republic*

(*Accepted 2 July 2003*)

SUMMARY

The original method for molecular typing of *E. coli* strains was developed using the polymorphism in chromosomal sequences of bacterial interspersed mosaic elements (BIMEs) detected by multiplex PCR and analysed by AFLP assay. The applicability of the method in the epidemiology of *E. coli* was tested on a group of 524 strains of human and veterinary origin. In the studied group 18 different genotypes were detected. Significant differences were found in the frequencies of the genotypes among various groups of strains, suggesting the method could be a promising tool in the epidemiology of *E. coli*.

INTRODUCTION

Over the past years highly precise and reproducible molecular genotyping methods have been implemented into routine epidemiology studies of infection diseases. Various methods have been used for the characterization of *E. coli* strains. The results of ribotyping, amplification fragment length polymorphism (AFLP) [1], pulse field gel electrophoresis (PFGE), analysis of genes encoding antibiotic resistance and virulence factors or plasmid profile analyses are valid for the epidemiology of *E. coli*. However, they are time-consuming, laborious and may lack adequate discriminatory power [2, 3]. Molecular epidemiological analyses based on randomly amplified polymorphic DNA [4] may not be sufficiently robust and reproducible [1, 5]. PCR-based methods in bacterial epidemiology are based on the analyses of inter-strain sequence polymorphism of amplified regions [6]. The method presented here uses multiplex PCR technology and permits simultaneous amplification of several genetically variable bacterial interspersed mosaic elements (BIMEs) of the *E. coli* genome [7, 8]. The epidemiological applicability of the method was tested on groups of 350 *E. coli* isolates from beef meat, milk and milk products and 174 *E. coli* isolates cultured from human faeces and urine.

MATERIALS AND METHODS

Bacterial strains

(1) A total of 350 veterinary *E. coli* isolates from various materials (milk, meat, milk products) and localizations (milk farms) were collected at the Veterinary Research Institute, Brno (Table 1).
(2) A total of 174 human *E. coli* isolates from different clinical materials (faeces, urine) were collected at the Institute of Microbiology, University Hospital, Brno (Table 1).

* Author for correspondence.

https://doi.org/10.1017/S095026880300133X Published online by Cambridge University Press

Table 1. *The frequencies of individual genotypes* (*A–S*) *in studied groups* (*all values are %*)

Genotype	Milk total (*n*=204)	Milk farm 1 (*n*=100)	Milk farm 2 (*n*=104)	Milk products (*n*=45)	Beef meat (*n*=101)	Veterinary strains – total (*n*=350)	Human stool (*n*=86)	Human urine (*n*=88)	Human strains – total (*n*=174)	Total (*n*=524)
A	56·37	41·00	71·15	28·88	23·76	43·43	2·33	18·18	10·34	32·44
B	28·43	39·00	18·27	22·22	7·92	21·71	22·09	4·55	13·22	18·89
C				2·22	29·70	8·85	13·95		6·89	8·21
D					2·97	0·86				0·66
E					2·97	0·86				0·60
F	12·75	18·00	7·69	15·55	20·79	15·43	9·30		4·60	11·83
G				24·44	2·97	4·00	15·12	2·27	8·62	5·53
H	2·45	2·00	2·88		2·97	2·28	5·81		2·87	2·48
I					2·97	0·86		9·09	4·60	2·10
J				2·22		0·29	6·98	59·09	33·33	11·26
K				4·44		0·57	1·16		0·57	0·60
L					1·98	0·57	9·30		4·60	1·91
M					0·99	0·29				0·20
N							1·16	1·14	1·15	0·40
O							3·49		1·72	0·60
P								1·14	0·57	0·20
R							1·16	4·55	2·87	0·95
S							8·14		4·02	1·34

Table 2. *The oligonucleotides used as PCR primers*

Primer pairs	Genes	Oligonucleotide sequences
mtlA	Mannitol permease	5′-CTGCTGGCAGGTCGTAAGTA-3′
mtlD	Mannitol-phosphate dehydrogenase	5′-TTACCGATAAAGCCACGACCGAT-3′
araA	Periplasmic protein	5′-CCATATCGTCAGCGGTCATGA-3′
araD	Maltose high-affinity receptor	5′-CTGCGCTGGAACGAAGTGT-3′
lamB	L-arabinose isomerase	5′-CAGCTTCGGTCGTGGCGACA-3′
malM	L-ribulose 5-phosphate 4-epimerase	5′-CGTAGTTAACATCGGCAAGGCT-3′

The strains were isolated on blood agars, selected on Chromocult agar and identified by API 20E system (bioMérieux, La Balme-les-Grottes, France).

Sample preparation

Suspensions of bacterial strains in distillated water at concentrations of approximately 10^6–10^8 c.f.u./ml were collected into 2 ml Eppendorf vials, heated for 15 min at 100 °C in a dry block and centrifuged for 10 min at 10 000 ***g*** at room temperature. Supernatants were directly added to the polymerase chain reaction (PCR) mixture.

PCR

The total 50 μl reactions contained 2 U DynaZyme polymerase (FinZyme, Finland), 5 μl of supplied 10× reaction buffer, 200 mM dNTP mixture (Roche, Prague, Czech Republic), 2% glycerol, additional $MgCl_2$ to a final concentration of 2 mM, 5 μg/ml Red Cresol (Sigma, Prague, Czech Republic), 5 μl of sample, primer pairs *mtlA* and *mtlD*, *lamB* and *malM* and *araA* and *araD* (Table 2) in concentrations of 12, 20 and 40 pmol of each per reaction, respectively. The reactions were performed in a PTC 200 thermocycler (MJ Research, IL, USA) under the following conditions: initial denaturation at 96 °C for 6 min; 55 cycles of 96 °C for 10 s, 60 °C for 15 s, 72 °C for 50 s; final elongation at 72 °C for 2 min. The amplification products were separated on 3% ethidium bromide-stained agarose gels at 3 V/cm, visualized by UV illumination at 312 nm and photographed by CCD camera. The lengths of electrophoretic bands were measured using G-Gen software (prepared in our laboratory, under authorization).

https://doi.org/10.1017/S095026880300133X Published online by Cambridge University Press

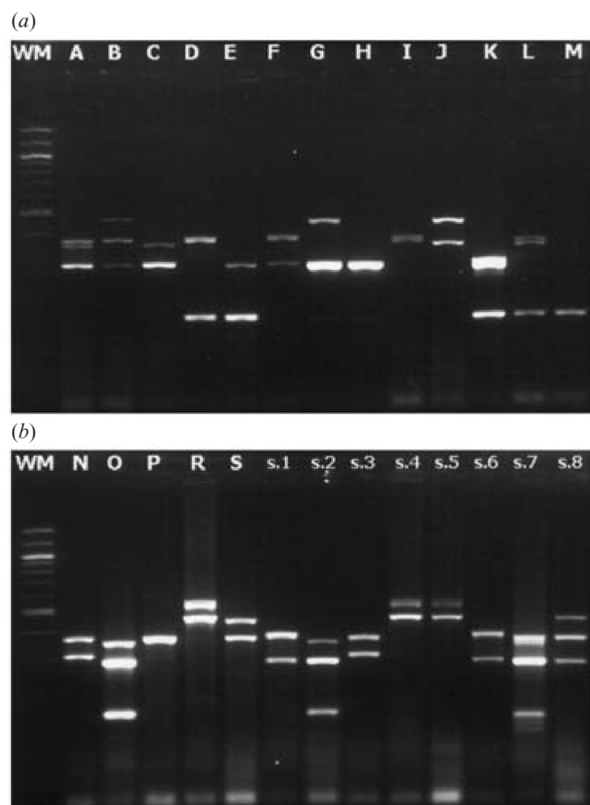

Fig. 1. The genotypes detected in the study. WM, Molecular weight marker (100 bp ladder, Takara, Otsu, Shiga, Japan), genotypes A–M (*a*) and genotypes N–S (*b*). s.1–s.8, Genotyped samples (various *E. coli* strains).

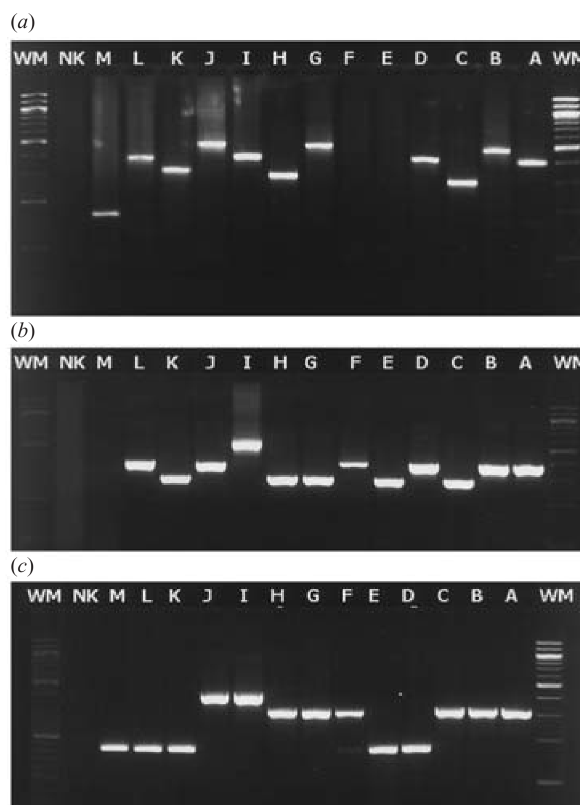

Fig. 2. The genotyping with individual primer pairs: (*a*) *araA* and *araD*; (*b*) *malM* and *lamB*; (*c*) *mtlA* and *mtlD*, illustrated on genotypes M, L, K, J, I, H, G, F, E, D, C, B, A. WM, Molecular weight marker (100 bp ladder, Takara), NK, negative control.

RESULTS

The method based on the combination of AFLP and multiplex PCR was developed. The patterns obtained after amplification were combinations of three different amplification products of variable sequences localized between the genes for *mtlA* and *mtlD*, *lamB* and *malM*, and *araA* and *araD*. The three primer pairs (Table 2) were designed by ClustalW software using genomic sequences of *E. coli* (GeneBank, accession nos. AP002568, AE005587, AE000438, AP002566, U00039, X06794, V01503, X51359, Z75440, Z75417, Z75416, Z75415, Z75414, Z75413, Z75412) to the conserved sequences of the 3′ and 5′ ends of these genes to amplify variable intergenic sequences. The lengths of these amplification products were variable due to strain-specific arrangements of non-coding intergenic sequences of BIMEs, which are located between relatively conserved genes (Table 2).

The amplification fragments were separated on agarose gels and photographed by CCD camera. The method reproducibly allowed the distinguishing of differences of no less than 10 bp in length of the amplification fragments using standard agarose gel electrophoresis. The combinations of three, two or one band (160–490 bp) were detected (Figs 1 and 2). All 350 veterinary and 174 human *E. coli* strains included in the study were successfully genotyped. In the studied group 18 different genotypes (A–S) were observed (Fig. 1, Table 3). The reproducibility of the method was tested on four randomly selected *E. coli* strains, which were cultured for 1 year on blood agar media. The same specific patterns after 10, 20, 30, 40 and 50 generations were detected.

From 100 twice tested different *E. coli* strains 99 repeatedly yielded the same genotypes. Only one of them required reclassifying from genotype F to B.

Thirteen genotypes (A–M) were detected only in ‘veterinary’ *E. coli* strains. Significant differences among various groups of strains cultured from beef meat, milk and milk products were found in the frequencies of A, B and F genotypes ($\chi^2_{0\cdot95}=82\cdot34$ for genotype A, $\chi^2_{0\cdot95}=31\cdot96$ for genotype B, $\chi^2_{0\cdot95}=27\cdot56$ for genotype F). The genotypes C and G

https://doi.org/10.1017/S095026880300133X Published online by Cambridge University Press

Table 3. *The genotypes detected in the study*

Genotype	Fragment lengths ± S.D.*	No. of observed fragments
A	291 ± 7, 359 ± 8, 379 ± 9	3
B	291 ± 8, 379 ± 11, 479 ± 10	3
C	291 ± 7, 359 ± 8	2
D	170 ± 8, 379 ± 10	2
E	170 ± 7, 291 ± 7	2
F	291 ± 6, 379 ± 10	2
G	291 ± 7, 479 ± 12	2
H	291 ± 8	1
I	359 ± 8, 379 ± 11	2
J	359 ± 2, 479 ± 12	2
K	170 ± 8, 291 ± 7, 308 ± 7	3
L	170 ± 5, 359 ± 8, 379 ± 11	3
M	170 ± 8	1
N	378 ± 4, 308 ± 3	2
O	170 ± 8, 291 ± 7, 359 ± 8	3
P	359 ± 8	1
R	479 ± 10, 528 ± 9, 554 ± 8	3
S	359 ± 7, 479 ± 12	2

* The fragments were detected on 3% agarose gel, their lengths (in bp) were measured and the S.D. values were counted by G-Gen software.

were frequently found in isolates from beef meat but not in milk. Significant differences ($P > 0{\cdot}95$) between two different sources of milk (farm 1 and farm 2) were observed again in the frequencies of genotypes A($\chi^2_{0{\cdot}95} = 20{\cdot}17$), B($\chi^2_{0{\cdot}95} = 10{\cdot}12$) and F($\chi^2_{0{\cdot}95} = 11{\cdot}81$) (Table 1).

Fifteen genotypes (A, B, C, F, G, H, I, J, K, L, N, O, P, R, S) were recognized in 'human' *E. coli* strains. The strains A and J were preferentially found in urine samples (isolated from patients with urinary-tract infection), whereas strains B, C, F, H and L were prevalent in stool samples. The genotype J was frequently and almost exclusively found in human strains (Table 1).

DISCUSSION

A simple, rapid, robust and reproducible method based on AFLP for molecular epidemiology of *E. coli* was developed. It was not possible to find any single region variable enough to obtain appropriate numbers of different strain-specific genotypes. Therefore we used the multiplex PCR method for simultaneous amplification of these three regions to combine their AFLP variability, the sources of which were BIMEs located between three pairs of genes [7]. For the purpose of epidemiological use we did not analyse the obtained polymorphisms in detail. The target sequences are sufficiently variable and optimal for agarose gel electrophoretical analyses [8]. The observed fragment lengths varied from 170 to 554 bp and their combinations produced 18 genotypes easy distinguishable on agarose gels. Theoretically, each primer pair should produce one fragment of specific length so all genotypes should be combined from three fragments after electrophoresis. However, some genotypes showed only one (genotypes H or M) or two (genotypes C, E) fragments. The PCR reactions with individual primer pairs produced a fragment of the same length so they overlapped each other in multiplex design forming double or multiple bands (Fig. 2). Moreover, several strains repeatedly did not yield any products with some primer pairs, probably due to mismatches in primer target sequences or integrated insertion sequence IS1397 [7].

The 'soft spot' of the presented method was a possible failure in the amplification of longer fragments. When PCR was slightly inhibited such fragments were sometimes unable to be detected or could only be recognized as weak signals. We believed that this was the reason why the one of 100 twice-tested different strains needed to be reclassified from genotype F to B after repeated amplification. Therefore, the reproducibility of the method depended on the efficiency of the reaction, although the quality of the template could also affect the results. From this point of view the increasing of PCR cycles to 55 proved to be beneficial. Nevertheless the lengthy passage of four *E. coli* strains proved the genotypes were stable even after 50 passages making the method useful even for long-term epidemiological analyses. The applicability of the method as an epidemiological tool was tested on the group of 350 *E. coli* strains cultured from beef meat, milk and milk products and on 174 *E. coli* strains cultured from human faeces and urine. Eighteen different genotypes (A–S) were found, some of which were preferentially detected in certain specimens (genotype J in urine, genotype C in human faeces and in cow milk, etc.). Significant differences were found in the frequencies of these genotypes among various groups of different origins (milk farms 1 and 2). From the data presented we concluded that this method can reliably detect and recognize various *E. coli* genotypes (Table 1).

There are many epidemiological methods based on detection and analysis of resistance genes, which can be transferred from cell to cell by bacterial conjugation or phage transduction. These horizontal gene

https://doi.org/10.1017/S095026880300133X Published online by Cambridge University Press

transfers could be enforced and accelerated by antibiotic therapy, which acts as specific selection mechanisms in some bacterial populations. Therefore these methods only follow up the epidemiology of resistance genes or plasmids, not of the bacteria themselves.

The method presented here is based on polymorphism detection in genetically variable BIMEs directly on bacterial chromosomes. Therefore it is independent on horizontal gene transfers and could only follow up vertical gene transfers through the lineages of *E. coli* generations.

The presented method, when tested on bigger groups of epidemiologically different sources, may be a promising tool for simple, rapid, robust and reproducible molecular epidemiology of *E. coli*.

ACKNOWLEDGEMENT

This work was supported by grant no. QC0196/200 of the National Agency for Research in Agriculture of the Czech Republic.

REFERENCES

1. Arnold C, Metherell L, Clewley JP, Stanley J. Predictive modeling of fluorescent AFLP: a new approach to the molecular epidemiology of *E. coli*. Res Microbiol 1999; **150**: 33–44.
2. Koort JM, Lukinmaa S, Rantala M, Unkila E, Siitonen A. Technical improvement to prevent DNA degradation of enteric pathogens in pulsed-field gel electrophoresis. J Clin Microbiol 2002; **40**: 3497–3498.
3. Michálek J, Collins RH, Durrani HP, Douek DC, Vitetta ES. Donor-derived alloreactive T cell clones and anti-leukemia T cell clones can be different. Blood 2002; **100**: 411b.
4. Cave H, Bingen E, Elion J, Denamur E. Differentiation of *Escherichia coli* strains using randomly amplified polymorphic DNA analysis. Res Microbiol 1994; **145**: 141–150.
5. Meunier JR, Grimont PAD. Factors affecting reproducibility of random polymorphic DNA fingerprinting. Res Microbiol 1993; **144**: 373–379.
6. Bachellier S, Saurin W, Perrin D, Hofnung M, Gilson E. Structural and functional diversity among bacterial interspersed mosaic elements (BIMEs). Mol Microbiol 1994; **12**: 61–70.
7. Bachellier S, Clement JM, Hofnung M, Gilson E. Bacterial interspersed mosaic elements (BIMEs) are a major source of sequence polymorphism in *Escherichia coli* intergenic regions including specific associations with a new insertion sequence. Genetics 1997; **145**: 551–562.
8. Bachellier S, Clement JM, Hofnung M. Short palindromic repetitive DNA elements in enterobacteria; a survey. Res Microbiol 1999; **150**: 627–639.

https://doi.org/10.1017/S095026880300133X Published online by Cambridge University Press