Fingerprinting of Bacterial Genomes by Amplification of DNA Fragments Surrounding Rare Restriction Sites

INTRODUCTION

DNA sequences that are characteristic of certain strains of a bacterial species can be very interesting from the biological and medical point of view. Some of them may determine drug resistance or contain plasmids or segments of DNA responsible for pathogenicity, such as "pathogenicity islands" (PAIs), integrons, genes responsible for synthesis of bacterial surface structure components, or they may encode restriction enzymes.

Other polymorphic DNA sequences that may have little or no influence on the pathogenicity or fitness of a bacterium include insertion sequences, plasmids, and prophages but can serve as markers for species/strain identification or evolutionary studies.

A variety of PCR-based methods for displaying DNA sequence polymorphism have been developed. Some of the methods such as RAPD (5) and AFLP (11) do not require prior knowledge of the DNA sequence. RAPD allows detection of DNA polymorphisms between strains of a species but does not exhibit high rates of reproducibility. AFLP permits selective amplification of restriction fragments from a total restriction enzyme digestion of a DNA and identification of DNA polymorphism within a species but requires the use of sequencing gels and usually labeled primers because of the quantity of simultaneously amplified DNA fragments. Suppression subtractive hybridization (SSH) was successfully applied to find DNA sequences characteristic for only certain strains of a bacterial species (3).

Here, we describe a method of finding differences in DNA sequence among bacterial species strains that is based on digestion of total bacterial DNA with two restriction enzymes differing in cleavage frequency, ligation with two different oligonucleotides, and suppression of PCR (6,8,10). PCR suppression allows the amplification of only a limited subset of DNA fragments, as only those with two different oligonucleotides ligated at the ends of complementary DNA strands are amplified in the PCR. The method does not require prior knowledge of the sequence of the analyzed DNA and generates a limited number of DNA fragments, whose band pattern on the gel differs between strains of a bacterial species. Furthermore, the DNA fragments can be easily analyzed on polyacrylamide gels stained with ethidium bromide.

The method reveals length polymorphism and occurrence of DNA fragments bordered by cleavage sites of restriction enzymes. We have implemented this method using a set of clinical E. coli and Klebsiella sp. strains and ATCC E. coli MG 1655 strain whose genome had been fully sequenced (2).
E. coli DNA was digested with combinations of two enzymes; BamHI (10 U/µL; Amersham Pharmacia Biotech, Piscataway, NJ, USA)/XbaI (15 U/µL; Amersham Pharmacia Biotech), or Nod (10 U/µL; Amersham Pharmacia Biotech)/BglII (10 U/µL; Roche Molecular Biochemicals, Mannheim, Germany), and DNA from Klebsiella and E. coli number 2 was digested with BglII/XbaI. Digestion reactions were performed under uniform conditions; 5 µL DNA solution, 2 µL 10× buffer corresponding to the enzyme combinations described above, respectively: buffer A for BamHI/XbaI (33 mM Tris-acetate, pH 7.9, 10 mM MgCl₂, 66 mM NaAc, 0.5 mM DTT; Roche Molecular Biochemicals) and buffer H for other enzyme combinations (50 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 100 mM KAc, 1 mM DTE; Roche Molecular Biochemicals), 2 µL each one of the pair of the enzymes were added and 9 µL water at 37°C for 2 h. To each of the digested samples, 25 µL 3 M NaAc, 180 µL TE buffer, 600 µL 96% ethanol, and 3 µL glycogen (20 mg/mL) (USB™; Amersham Pharmacia Biotech) were added, samples were vortex mixed and incubated on dry ice for 10 min, then centrifuged at 12,000 × g for 10 min. Ethanol was removed, and the pellets were washed with 200 µL 70% ethanol, then centrifuged for 3 min at 12,000 × g, and again ethanol was removed, and the pellets were dried for 15 min at 20°C.

Ligation

Adapters were assembled from two oligonucleotides (Table 1). Oligonucleotides were dissolved in ligation buffer (66 mM Tris-HCl, pH 8.5, 6.6 mM MgCl₂, 10 mM DTT, 66 µM ATP; Amersham Pharmacia Biotech) to concentration of 20 pmol/µL (each oligonucleotide), then heated at 90°C in a water bath for 2 min and subsequently left at room temperature for 10 min to anneal. Appropriate adapters were ligated to the corresponding cohesive ends (Table 1); dry pellets of DNA digests were dissolved in a solution consisting of 2 µL 10× ligation buffer, 1 µL (20 pmol each adapter) of the solution of the adapters corresponding to the cohesive ends left by the enzymes used for the prior digestion of the sample, 1 µL 1 U/µL T4 DNA ligase (USB; Amersham Pharmacia Biotech), water to 20 µL at 16°C for 2.5 h or overnight. After ligation, 80 µL TE buffer were added, and DNA was isopropanol precipitated (1). The samples were dissolved in 20 µL TE buffer.

Table 1. Adapters and Corresponding Restriction Sites

<table>
<thead>
<tr>
<th>Adapter Restriction site</th>
<th>Adapter</th>
<th>Corresponding Restriction site</th>
</tr>
</thead>
<tbody>
<tr>
<td>5′-CCTCATCCACCAAGCTGAC-3′</td>
<td>POWIEBIS</td>
<td>XbaI †CTAGA</td>
</tr>
<tr>
<td>3′-TCGACGCTGATC-5′</td>
<td>XBALIG</td>
<td></td>
</tr>
<tr>
<td>5′-CCTCATCCACCAAGCTGAC-3′</td>
<td>POWIEBIS</td>
<td>NotI GC↓GGCAGC</td>
</tr>
<tr>
<td>3′-TCGACGCTGATC-5′</td>
<td>NOTLIG</td>
<td></td>
</tr>
<tr>
<td>5′-GGATGGTAGAAGAACGCTGAC-3′</td>
<td>BOXG</td>
<td>BglIII A↓GATCT</td>
</tr>
<tr>
<td>3′-CCTACCATCTGTCTCTTCTTGCAGCTGAGC-5′</td>
<td>BOXD</td>
<td>BamHI G↓GATCC</td>
</tr>
</tbody>
</table>

Adapters are double stranded but only their constant parts (bold) are ligated, helper parts serve just to create dsDNA fragment and protruding 5′ end.

All the oligonucleotides are not phosphorylated. Helper oligonucleotide anneals to the constant part and together form a fragment of dsDNA with a 5′ protruding end identical with the protruding end left by the corresponding restriction endonucleases. Helper oligonucleotides NOTLIG, XBALIG and BOXD are not phosphorylated and can not be ligated to the 3′ end of the restriction DNA fragment. The underlined regions of BOXD and BOXG are not complementary. Non-complementary region prevents formation of double stranded BOXD and BOXG particle stable under filling-in conditions (Figure 1 step 3). PCR primers: POWIEBIS and BOXPOWIE.

MATERIALS AND METHODS

Bacterial Strains

E. coli MG 1655 ATCC (ATCC, Manassas, VA, USA), and six E. coli strains of clinical origin were used. Strains 2, 4, and 5 contained several plasmids, two strains of Klebsiella oxytoca and four of Klebsiella pneumoniae, also of clinical origin, were used.

Isolation and Digestion of Bacterial DNA

DNAs from 3 mL overnight cultures of bacterial strains were isolated (1). Dry DNA pellets were dissolved in 150 µL TE buffer.

DNA quantities in the samples were estimated by electrophoresis of 2 µL DNA solution on 0.8% agarose (Sigma, St. Louis, MO, USA) gels with TAE buffer (9) run together with samples with known amounts of DNA and subsequent ethidium bromide staining (Sigma-Aldrich Chemie GmbH, Steinheim, Germany) in 0.5 mg/L solution for 10–15 min. Images of the gels were made using a White/Ultraviolet Transilluminator (UVP, San Gabriel, CA, USA).

The DNA concentration range was from about 100 ng/µL to several hundred ng/µL.

E. coli DNA was digested with about 100 ng/µL using a White/Ultraviolet Transilluminator (UVP, San Gabriel, CA, USA). with known amounts of DNA and subsequently ethidium bromide staining (Sigma-Aldrich Chemie GmbH, Steinheim, Germany) in 0.5 mg/L solution for 10–15 min.
PCR

The reaction was performed in 500 µL Eppendorf tubes in a 50-µL reaction mixture containing 5 µL 10× PCR buffer (Expand™ Long Template PCR System; Roche Molecular Biochemicals buffer 1, 1.75 mM MgCl₂), primers: 50 pM each BOXPOWIE and POWIEBIS (5′-GGATGGTAGACGAAGGAACGC-3′, 5′-CCTTCATCCACCAACGTGAC-3′; TIB Molbiol, Poznan, Poland, and Genset Oligos, Paris, France), 1 µL ligation solution, 100 µM each dNTPs, 15 µL betaine (5 M solution; Sigma), 2.5 U Taq DNA polymerase, and water to 50 µL. All the components of the PCR solution, except Taq DNA polymerase, were vortex mixed in a 500-µL PCR tube. Thermal cycling was performed in an Eppendorf thermal cycler (Mastercycler® 5330; Eppendorf Netheler Hinz GmbH, Hamburg, Germany), as follows: at 68°C for 2 min and held constantly until Taq DNA polymerase was added, at 68°C for 5 min for filling in the ends of the DNA fragments, followed by denaturation at 94°C for 90 s, then 19 cycles of 94°C for 30 s, 60°C for 30 s, 68°C for 1 min, followed by 5 min at 68°C.

PCR products (10 µL of 50 µL) were electrophoresed on 6% polyacrylamide gels [1:60 acrylamide (J.T. Baker, Philipsburg, NJ, USA)/N,N’-METHYLENE bis-acrylamide (Sigma-Aldrich Chemie GmbH, Steinheim, Germany)] with TAE buffer, stained in ethidium bromide (Sigma-Aldrich Chemie GmbH) at 0.5 mg/L aqueous solution for 10–15 min; images of the gels were made using a White/Ultraviolet Transiluminator (UVP).

A PCR product obtained from DNA of E. coli MG 1655 (ATCC) treated with a combination of enzymes BamHI/XbaI using the previously described procedure was electrophoresed on a 6% polyacrylamide gel, the DNA band which co-migrates with the 489-bp DNA fragment of the marker was excised from the gel. DNA was eluted (7). The DNA fragment was sequenced on Open Gene System Long-Read Tower™ (Visible Genetics, Toronto, Canada), using BOXPOWIE primer, and the Thermo Sequenase™ Cy™5.5 dye terminator cycle sequencing kit (Amersham Pharmacia Biotech). The obtained sequence was consistent with the one predicted by analysis of the E. coli MG 1655 ATCC DNA sequence with Vector NTI 4.0 Deluxe 32-bit software (Informax, Bethesda, MD, USA). The predicted length of the fragment was 411 bp, but the amplified fragment contained additional sequences derived from the ligated adapters.

RESULTS AND DISCUSSION

Outline of the Method

The PCR suppression phenomenon (6,8,10) is the basis for obtaining limited representation of the DNA fragments that form the bacterial genome. Our proposal is to amplify exclusively DNA fragments surrounding relatively rare nucleotide sequences (e.g., rare restriction sites). The outline of the method is shown in Figure 1. A DNA preparation
to be characterized is digested with two restriction enzymes, rare and frequent cutters. Three kinds of DNA fragments—abundant (A), sporadic (S), and limited (L)—arise that are formed after digestion with frequent, rare, and both cutters at the same time, respectively. The mixture of DNA fragments is ligated with two different synthetic adapters (Figure 1). All 5’ ends of the most abundant DNA fragments produced by digestion with a frequent cutter (A fragments) are modified by joining the same synthetic oligonucleotide (BOXG). Similarly, S fragments generated by digestion with a rare cutter are modified by ligation of oligonucleotide POWIEBIS to both 5’ ends of each dsDNA fragment. After filling in of the modifying oligonucleotides joined to the 5’ ends with DNA polymerase all single-stranded A and S DNA fragments have complementary sequences on their 5’ and 3’ ends, and, because of that, the proper usage of suppression PCR (SP PCR) during amplification of the genomic fragment mixture should eliminate the most and least abundant DNA fragments from the mixture, fragments A and S, respectively. However, L fragments arising after digestion with rare and frequent restriction enzymes at the same time and consequently joined with two distinct modifying oligonucleotides, POWIEBIS and BOXG, are amplified exponentially. After filling in of the modifying oligonucleotides joined to the 5’ ends with DNA polymerase, single-stranded L fragments do not have complementary sequences on their 5’ and 3’ ends and consequently are not susceptible to SP PCR. Furthermore, newly synthesized 3’-terminal sequences contain the primer binding portion: for primer POWIEBIS on one strand and for BOXPOWIE on the other.

Calibration of the Method

In principle, for bacterial genomes of known nucleotide sequences, the method should produce predictable results. To verify this point, we have analyzed the nucleotide sequence of *E. coli*.
K-12 MG1655 using Vector NTI 4.0 Delux 32-bit software to check the number and sizes of DNA fragments that should be obtained after digestion with two pairs of restriction nucleases used: XbaI (rare cutter), BglII (frequent cutter) and NotI (rare cutter), and BglII (frequent cutter). The results of the analysis were compared with experimental data obtained after the application of our method to the total DNA preparation isolated from E. coli MG1655. It appears that, for both pairs of restriction nucleases, in the range of up to about 1000 bp, the electrophoretic patterns, number, approximate length, and relative mobility of the amplified fragments agree with the computer analysis. After treatment of E. coli MG1655 DNA with BglII and XbaI restriction nucleases (Figure 2, left side), the following DNA fragments should be visible: 364, 464, 541, 755, and 1004 bp (the length of both adapters equal to 53 nucleotides was added to that of the restriction fragments). The DNA fragment, which theoretically should migrate as a 464-bp fragment, co-migrates with the 489-bp DNA fragment of the marker (Figure 2). However, sequence analysis of the DNA eluted from the gel strip containing this DNA fragment confirmed its identity as the 464-bp long fragment that was predicted by computer analysis.

For the NotI and BglII pair of enzymes, the amplified fragments should be 182, 669, 693, 706, and 876 bp in length.

**Fingerprinting of Bacterial Strains**

We applied our method to analyze total DNA from a number of E. coli strains, two strains of K. pneumoniae and four of K. oxytoca. In this report, we present results for 7 of the 17 analyzed E. coli strains (Figure 2). Six of the presented E. coli strains and all of the Klebsiella strains were of clinical origin, and the seventh E. coli strain presented was MG1655 (ATCC), whose genome had been completely sequenced (2).

Only two E. coli strains (nos. 3 and 5) of the 17 analyzed have identical DNA band patterns after electrophoresis of PCR products (Figure 2). This observation is valid for both pairs of restriction nucleases used.

For all analyzed E. coli strains under analysis, in all the PCR products from the BglII/XbaI digest, one DNA band about 489 bp long is present (Figure 2). Nucleotide sequencing of this DNA fragment revealed that it is identical with part of E. coli MG1655 genome (GenBank accession no. NC_000913 positions 921487-921898). The internal part of the fragment encodes a cold shock protein (CSPD), whose sequence is highly conserved (protein ID no. AACT73967.1). For the BglII/NotI digest, two DNA bands are common among all the analyzed E. coli strains (Figure 2).

As expected, the DNA band patterns depend on more than the presence or absence of an extrachromosomal DNA: (i) E. coli strains 6 and MG1655 (ATCC) do not bear plasmids and nevertheless have different DNA band patterns; (ii) strain nos. 3 and 5 have the same DNA band pattern of the PCR products (Figure 2), although strain no. 5 hosts several plasmids and strain no. 3 bears only one of high molecular weight; and (iii) strain no. 4 bears the same plasmids as strain no. 5 and has a different DNA band pattern of PCR products (Figure 2).

The high differentiation power of the method is shown on clinical strains of Klebsiella. In this case, the XbaI/BglII pair of restriction nucleases was used at the first step of total bacterial DNA treatment. Figure 3 shows a high degree of diversity of the analyzed strains, but there is at least one DNA band whose size is identical in all the analyzed PCR products from Klebsiella and has about 1000 bp (Figure 3).

The same combination of restriction nucleases, when applied to the E. coli strain no. 2 total DNA preparation, resulted in a higher number of PCR products (results not shown), within the analyzed DNA fragment length range, compared to other pairs of restriction nucleases used. Although computer analysis does not reveal a significant difference in the amount of restriction fragments, within the analyzed size range, generated by each of three combinations of restriction nucleases combinations applied to analyze the E. coli K-12 MG1655 genome. It might sug-
suggest that the BglII/XbaI pair of nucleases is the most efficient in differentiation of the analyzed bacteria.

Reproducibility of the Method

Identical results were obtained in independent experiments performed with DNA isolated from different cultures of one strain. Results obtained in PCRs do not depend on the thermostable DNA polymerase used (results not shown) or on the thermal cycler used (results not shown). The use of betaine improves the PCR efficiency, especially for high molecular weight DNA fragments. The relative amount of PCR products within an amplified sample is also characteristic for the bacterial strain and the combination of restriction nucleases applied. Some DNA fragments were amplified more efficiently (Figure 2).

Choice of Enzymes and Oligonucleotides

Whether a restriction endonuclease is a frequent or rare cutter depends on GC content and ultimately on short oligonucleotide frequencies in the analyzed genome. The restriction endonucleases NotI and XbaI are rare cutters for genomes of Enterobacteriaceae, as presented in this paper, but, in some genomes with different GC contents, they could be frequent cutters. For the Mycobacterium tuberculosis genome (4), (GenBank accession no. NC_000962), NotI restriction nuclease is a frequent cutter (1216 sites vs. 19 for the E. coli genome). Restriction nuclease XbaI is an intermediate cutter for the M. tuberculosis genome (109 sites vs. 39 for the E. coli genome).

There is one 119-bp DNA fragment bordered by sites of XbaI restriction nuclease in the genome of E. coli K-12 MG1655 (2). We have not observed amplification of this fragment. The oligonucleotide ligated to the protruding ends of DNA fragments digested with the XbaI endonuclease is also used as a PCR primer (POWIEBIS). After ligation of the oligonucleotides and filling in the ends of those DNA fragments, each DNA strand is terminated with complementary sequences five bases longer than the sequence of POWIEBIS. The five bases derive from the 5' end of the restriction fragment, and the length of the complementary sequences in the terminal parts of the DNA fragments is obviously sufficient to cause efficient suppression of PCR amplification, at least for DNA fragments of relatively small size, such as the 119-bp XbaI-XbaI fragment of E. coli K-12 MG1655. This fact should be taken into account when designing oligonucleotides for experiments involving ligation-mediated PCR.

It is probable that a five-base difference in length of PCR primer and complementary sequences would not cause efficient SP PCR for relatively long DNA fragments; therefore, those fragments could be amplified and contribute to the differentiation of the analyzed strains. The problem of SP PCR efficiency has been comprehensively discussed (10).

CONCLUSION

The advantages of this method include the following: (i) the method does not require prior knowledge of an analyzed sequence; (ii) results can be easily analyzed even on polyacrylamide gels stained with ethidium bromide; (iii) one set of adapters and enzymes can be applied to analyze DNA from diverse species of bacteria; (iv) PCR products can be directly isolated from the polyacrylamide gel and subsequently sequenced; and (v) the method can be set up (calibrated) on genomic DNA of an organism whose genome nucleotide sequence is already known.

ACKNOWLEDGMENTS

This work was supported by grant no. 4 PO5D 020 13 from The State Committee for Scientific Research (Poland). We thank Professor Danuta Dzierżanowska from The Childrens Memorial Health Institute, Warsaw, for her gift of bacterial strains.

REFERENCES


Received 20 November 2000; accepted 27 April 2001.

Address correspondence to:
Dr. Andrzej Plucienniczak
Staryosciniska 5
02-516 Warsaw, Poland
e-mail: ap@iba.waw.pl