

REVIEW ARTICLE

## Molecular methods for bacterial genotyping and analyzed gene regions

İbrahim Halil Yıldırım<sup>1</sup>, Seval Cing Yıldırım<sup>2</sup>, Nadir Koçak<sup>3</sup>

<sup>1</sup>Dicle University, Faculty of Veterinary Department of Genetics, Diyarbakır, Turkey

<sup>2</sup>İnönü University, Faculty of Science, Department of Biology, Malatya, Turkey

<sup>3</sup>Selçuk University, Selçuklu Medical Faculty, Department of Medical Genetics, Konya, Turkey

### ABSTRACT

Bacterial strain typing is an important process for diagnosis, treatment and epidemiological investigations. Current bacterial strain typing methods may be classified into two main categories: phenotyping and genotyping. Phenotypic characters are the reflection of genetic contents. Genotyping, which refers discrimination of bacterial strains based on their genetic content, has recently become widely used for bacterial strain typing. The methods already used in genotyping of bacteria are quite different from each other. In this review we tried to summarize the basic principles of DNA-based methods used in genotyping of bacteria and describe some important DNA regions that are used in genotyping of bacteria. *J Microbiol Infect Dis* 2011;1(1):42-46.

**Key words:** Bacterial genotyping, pulsed-field gel electrophoresis, internal transcribed region (ITS)

### Bakteri genotiplendirmesinde kullanılan moleküler yöntemler ve incelenen gen bölgeleri

#### ÖZET

Bakteri tiplendirilmesi tanı, tedavi ve epidemiyolojik araştırmalar için önemli bir süreçtir. Mevcut bakteriyel suş tiplendirme yöntemleri iki ana kategoride sınıflandırılabilir: fenotipleme ve genotipleme. Fenotipik karakterler genetik içeriğin yansımasıdır. Bakterilerin genetik içeriklerine bağlı olarak sınıflandırılması olan genotiplendirme son dönemlerde bakteri tiplendirmesinde sıklıkla kullanılmaktadır. Bakteri genotiplendirmesinde kullanılan yöntemler birbirlerinden oldukça farklıdır. Bu derleme ile bakteri genotiplendirmesinde kullanılan DNA temelli yöntemlerin temel prensiplerini özetleme ve bakteri genotiplendirmesinde kullanılan DNA bölgelerini açıklamaya çalıştık.

**Anahtar kelimeler:** Bakteriyel genotiplendirme, pulsed-field jel elektroforezi, internal transcribed region (ITS)

#### INTRODUCTION

Identifying of bacteria at the strain level is especially important due to the some features of bacteria that are also challenge to human health, including increased virulence and transmissibility, resistance to antibiotics, expanding host spectra and possibility of usage for bioterrorism after genetic manipulations.<sup>1</sup> There are two distinct methods used in bacterial identification: phenotyping and genotyping.<sup>2</sup> In phenotyping method; identifying of strains are based on phenotypic characters including morphology of colonies in various culture media, biochemical tests, serology, pathogenicity and antibiotic susceptibility.<sup>1,3</sup> Discriminating of closely related strains by these methods is not enough and characterization of cells on morphology, staining and metabolic traits

for unambiguous identification can take days to weeks.<sup>1,3</sup> Phenotypic characters such as the pathogenicity, host specificity, antibiotic resistance, virulence and geographic distribution of bacteria is closely related with their genetic diversity.<sup>1</sup> Various genetic methods have been developed for genotyping of bacteria, since the late 1980.<sup>4</sup> These methods have become frequently used in bacterial identification due to their high resolution.<sup>1</sup> The identified genetic profile of any bacteria by a specific genotyping method can be as unique as fingerprint.<sup>1</sup>

Current bacterial genotyping methods can be categorized into three categories: DNA banding pattern, DNA sequencing, and DNA hybridization.<sup>5,6</sup> In DNA banding method, DNA bands can be directly generated by digestion of

**Correspondence:** Assoc. Prof. İbrahim Halil Yıldırım

Dicle University, Faculty of Veterinary Department of Genetics, Diyarbakır, Turkey Email: ihyildirim@yahoo.com

Received: 19.05.2011, Accepted: 27.06.2011

Copyright © Journal of Microbiology and Infectious Diseases 2011, All rights reserved

Restriction Endonucleases (REs) or by amplification of known or unknown region of the genome or by a combination of amplification and digestion with restriction enzymes.<sup>1</sup> In DNA sequencing based method, discrimination among the bacterial strains performed after the determination and comparing of a known gene sequence.<sup>1</sup> In DNA hybridization based methods, discrimination of bacteria are carried out by analyzing the hybridization of known probes.<sup>6</sup> DNA macroarray and microarray systems have also been developed to get accurate and faster results in description of bacteria.<sup>1,4</sup>

This review briefly summarizes the basic principles of DNA-based methods, used in genotyping of bacteria, and describes some important DNA regions that were used in genotyping of bacteria.

### **DNA banding based methods**

DNA banding-pattern based methods classify the bacteria according to the size of fragments that generated by PCR amplification or digestion of genomic DNA by restriction endonuclease or combination of both digestion and amplification.<sup>1,6</sup> Identifying of generated bands can be determined by conventional agarose gel electrophoresis or automated capillary electrophoresis systems.

## **RESTRICTION ENZYME BASED METHODS**

### **Pulsed-field gel electrophoresis (PFGE)**

DNA molecules can be separated in conventional constant electrical field depending on the size of DNA. DNA fragments which are larger than 20 kb show the same mobility through a gel and move together in a size-independent manner under the constant electric current.<sup>1,7</sup> Separating larger DNA molecules can be achieved by applying alternating electric fields at different angles.<sup>2</sup> This method was described in 1984 firstly and became known as Pulsed Field Gel Electrophoresis (PFGE).<sup>2,7</sup> PFGE is an alternative method of restriction digest involves the use of REs with uncommon recognition motifs to generate large DNA fragments.<sup>7</sup> The banding patterns obtained from a group of strains reflect the DNA polymorphism at the RE recognition sites.<sup>1</sup> Although PFGE widely used in epidemiologic and environmental study, it has several limitations including DNA density, amounts of agarose in the gel, applied voltage and gel temperature.<sup>1</sup> Hence, identifying criteria

of bacteria by PFGE and PFGE protocols require to be standardized. Tenover and colleagues proposed the bacteria that have the same PFGE profile should be considered as belonging to the same strain.<sup>2</sup> They also proposed, isolates that differ by a single genetic event, which is reflected as a difference of one to three bands should be considered 'closely related', and isolates differing by four to six bands, likely representing two independent genetic events, should be considered 'possibly related', bacterial isolates containing six or more band differences should be considered 'unrelated'.<sup>2</sup>

### **Restriction fragment length polymorphism (RFLP)**

RFLP analysis is based on the measurement of fragments that are resulted from the digestion of genomic DNA.<sup>9,10</sup> Unlike PGFE, hundreds of short restriction fragments may produce due to digestion of genomic DNA with frequently cutting Res.<sup>1,2</sup> Because of difficulties in analyzing of many bands, hybridization with known DNA probs are often used in this technique.<sup>1</sup> For example, using of probes derived from IS6110, insertion element is gold standard method for typing *Mycobacterium tuberculosis* complex.<sup>11</sup>

rRNA probes can also be used for RFLP analysis.<sup>12</sup> In this method, called ribotyping, conserved regions of 16S and 23S rRNA genes were used for probing the different RFLP banding patterns.<sup>12</sup> This method allows determining the DNA sequence of fragments, because rRNA operons are universal.<sup>1</sup>

## **POST POLYMERASE-CHAIN REACTION (PCR) BANDING METHODS**

### **Random amplification of polymorphic DNA (RAPD)**

The RAPD method, also known as AP-PCR, based on the random amplification of unknown genomic regions by using single short primers.<sup>13</sup> Unlike classical PCR analysis, the genomic region that will be amplified is not known and amplification depends on the positions that are complementary to the primer sequences.<sup>13</sup> If primers anneal too far apart or a mutation has occurred at a site that was previously complementary to the 10-mer identical primers, amplification may not be performed.<sup>2</sup> The different-sized amplicons produced at multiple loci by RAPD-PCR may be

observed in different band patterns on agarose gel and bacteria can be genotyped depending on these band patterns.

### **PCR-RFLP method**

This method based on the digestion and separation of fragments in agarose gel electrophoresis after amplification of a specific locus with specific primers.<sup>2</sup> The difference of this method from direct RFLP method is the limiting of the interested DNA region. This method has been used widely for typing of a variety of bacteria and discriminatory power of this technique may be enhanced by multilocus-based analysis.<sup>1</sup>

### **Amplified fragment length polymorphism (AFLP)**

This technique based on the digestion of genomic DNA with two restriction enzymes and ligation of restriction fragments with end-specific adaptors and then amplification of fragments by complementary primers to adapter sequences.<sup>14</sup> As in other DNA banding pattern based methods, identifying of band patterns can be detected by conventional gel electrophoresis.<sup>14</sup> In addition, DNA sequence of fragments can be determined by using the primers that were designed against to the adapter sequences.<sup>14</sup> Restrictive side of this sensitive and high discriminatory method is that the DNA template should not be contaminated by various DNAs.<sup>1</sup>

### **PCR AND BANDS OF REPETITIVE DNA REGIONS**

#### **REP-PCR**

There are a series of repetitive DNA sequences that are dispersed in multiple copies throughout the bacterial genomes.<sup>15</sup> The functions of these interspersed repetitive DNA elements are still unknown.<sup>15</sup> Repetitive sequences can be categorized into three families: the 35-40 bp repetitive extragenic palindromic (REP) sequence,<sup>16</sup> the 124-127 bp enterobacterial repetitive intergenic consensus (ERIC) sequence,<sup>17</sup> and the 154 bp BOX element sequences.<sup>18</sup> The consensus sequences of those elements can be used for amplifying the DNA fragments between these repetitive elements.<sup>1</sup> The band patterns obtained from these amplifications are useful for DNA fingerprinting of bacteria.<sup>1</sup> Detection of bands can be performed by agarose gel electrophoresis or cap-

illary electrophoresis.<sup>1</sup> Reproducibility of the REP-PCR is much higher than the RAPD-PCR due to the used specific primers for amplification.<sup>1</sup>

### **Multiple locus variable number tandem repeat analysis**

MLVA is a genotyping tool that provides data based on the number of repetitive sequences found on the bacterial chromosomes.<sup>19</sup> Variable numbers of tandem repeats (VNTR) are the consecutive repeats that are dispersed in multi copies on the bacterial genome.<sup>19</sup> VNTRs can be found in noncoding regions as well as in genes and the number of tandem repeats may vary between strains.<sup>1,19</sup> The number of repeats can be determined by using the primers that are complementary to the well conserved regions of flanking the tandem repeats. Discrimination performed by comparing the PCR products for determining the relative degree of bacteria.<sup>19</sup>

### **DNA SEQUENCING BASED METHODS**

#### **Sanger method**

The Sanger method is also known as dideoxy or chain termination method, based on the synthesizing of DNA chain through the use of dideoxynucleotides that interrupt the elongation step of DNA amplification.<sup>20</sup> When DNA polymerase enzyme plugged a nucleotide without 3' hydroxyl group to the chain, elongation will be terminated.<sup>20</sup> The dideoxy nucleotide terminated the chain can be determined by separating the PCR products in acryl amide gel electrophoresis.<sup>20</sup>

#### **Pyrosequencing**

Unlike Sanger method, pyrosequencing is not an electrophoretic method. This method based on the real time detection of released pyrophosphates during DNA chain elongation.<sup>21</sup> Like classical PCR, Pyrosequencing method also requires primers for chain elongation. Unlike standard PCR, pyrosequencing require ATP sulfurylase, Luciferase, Apyrase and Adenosine phosphosulfate.<sup>21</sup> The four enzymes in to the pyrosequencing system are the Klenow fragment of DNA Polymerase I, Luciferase, ATP sulfurylase and Apyrase.<sup>21</sup> The reaction mixture also contains adenosine phosphosulfate, D-luciferin, DNA template and an annealed primer to be used as starting material. Klenow Polymerase incorporates the added dNTPs into the growing DNA strand and pyrophosphate

is released. ATP sulfurylase converts the pyrophosphate into ATP and Luciferase produce light by using the generated ATP. The four nucleotides are added one at a time and a camera detects the produced light as evidence of the incorporated nucleotide.<sup>21</sup>

### DNA hybridization based methods

Hybridization based this method, require a target DNA and a fluorescently labeled DNA fragment that is complementary to target DNA, called as prob. Hybridization technique can be applied to hundreds or tens of thousands of DNA fragments or oligonucleotides arrayed on a substrate. The arrays can be classified to microarray and macroarray according to the size and spots on the supports. These hybridization based array techniques are frequently used in the screening of mutation and bacterial genotyping studies.<sup>4</sup>

## THE GENE REGIONS USED IN BACTERIAL GENOTYPING

### The 16S rRNA gene

rRNA genes are the essential genes for the survival of all organisms due to their role in protein synthesis.<sup>1</sup> The 16S rRNA gene is about 1500 bp long and it is a composed of well conserved 10 regions and 10 divergent regions.<sup>4</sup> There is a constant mutation rate of about 1% per 50 years in the divergent regions of the 16S rRNA gene. Because of this polymorphic situation, rRNA gene sequences are used for more than 20 years in phylogenetic examinations.<sup>4</sup> In some cases, foreign DNA sequences, that are named intervening sequences of about 140 bp long can be found in the 16S rRNA gene so, if a 16S rRNA gene is larger than the usual size of about 1500 bp, it should be identified the possible IVS sequences in the 16S rRNA gene.

### 16S-23S rRNA ITS region

The 16S, 23S and 5S rRNA genes of prokaryotic microorganisms are found in the same genetic locus and they are separated by noncoding regions called internal transcribed spacer (ITS).<sup>1,22</sup> ITSs are special regions exhibited high rate of polymorphisms and large degree of length variations at the genus and species levels and therefore they are useful for identifying and subtyping of bacteria.<sup>22</sup> ITS regions can be easily amplified with the primers designed for complementary to the con-

served regions of 16S and 23S rRNA genes.<sup>1</sup> The ITS regions are more informative than the 16S rRNA analysis particularly in strain typing.<sup>1</sup>

### Housekeeping genes

Housekeeping genes are the genes required for basic functions of cell or organisms and they are essential for survival.<sup>23</sup> These genes can be used for discrimination of closely related strains that cannot be separated by 16S rRNA analysis. The *rpoB* gene which is a subunit of RNA polymerase enzyme widely used in bacterial genotyping. Some other housekeeping genes such as *hsp65*, *gyrB*, *slpA* genes are often used in bacterial genotyping.<sup>1,4,7</sup>

### Conclusion

The methods used for classical typing systems are based on the phenotypic characteristics of bacteria that are resulted of genomic formation. Genotyping can be evaluated a more detailed analyzing of phenotypic characters. ITS region or 16S rRNA gene can be more discriminative especially in closely related strains. In our assessment, it will be better to study more than one genomic region in identifying the closely related species or strains.

### REFERENCES

1. Li W, Raoult D, Fournier P. Bacterial strain typing in the genomic era. *FEMS Microbiol Rev* 2009;33:892-916.
2. Tenover FC, Arbeit RD, Goering RV. How to select and interpret molecular strain typing methods for epidemiological studies of bacterial infections: a review for healthcare epidemiologists. *Infect Control Hosp Epidemiol* 199;18:426-439.
3. Jackson GW, mcNichols RJ, Fox GE, Wilson RC. Bacterial genotyping by 16S rRNA mass cataloging. *BMC Bioinformatics* 2006;7:321 doi:10.1186/147-2105/7/321.
4. Cai HY, Archambault M, Gyles CL, Prescott JF. Molecular genetic methods in veterinary clinical bacteriology laboratory: current usage and future applications. *Animal Health Res Rev* 2003;4:73-93.
5. Taha MK, Olcen P. Molecular genetic methods in diagnosis and direct characterization of acute bacterial central nervous system infections. *APMIS* 2004;112:753-770.
6. Versalovic J, Lupski JR. Molecular detection and genotyping of pathogens: more accurate and rapid answers. *Trends Microbiol* 2002;10:S15-S21.
7. Khrosravi AD, Seghatoleslami S. Genotyping and identification of mycobacteria by fingerprinting techniques. *Jundishapur J Microbiol* 2009;2:81-91.
8. Tenover FC, Arbeit RD, Goering RV, Mickelsen PA, Murray BE, Persing DH, Swaminathan B. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J Clin Microbiol* 1995;33:2233-2239.



9. Panneerchelvam S, Norazmi MN. Forensic DNA profiling and database. *Malaysian J Med Sci* 2003;10:20-26.
10. Todd R, Donoff RB, Kim Y, Wong DT. From the chromosome to DNA: restriction fragment length polymorphism analysis and its clinical applications. *J Oral Maxil Surg* 2001;59:660-667.
11. Thierry D, Matsiota-Bernard P, Pitsouni E, Costopoulos C, Guesdon JL. Use of the insertion element IS6110 for DNA fingerprinting of *Mycobacterium tuberculosis* isolates presenting various profiles of drug susceptibility. *FEMS Immunol Med Microbiol* 1993;6:287-297.
12. Bingen EH, Denamur E, Elion J. Use of ribotyping in epidemiological surveillance of nosocomial outbreaks. *Clin Microbiol Rev* 1994;7:311-327.
13. Bart A, Schuurman IG, Achtman M, Caugant DA, Danker J, van der Ende A. Randomly amplified polymorphic DNA genotyping of serogroup A meningococci yields results similar to those obtained by multilocus enzyme electrophoresis and reveals new genotypes. *J Clin Microbiol* 1998;36:1746-1749.
14. Blears MJ, De Grandis SA, Lee H, Trevros JT. Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. *J Indust Microbiol&Biotechnology* 1998;21:99-114.
15. Vesolovic J, Koeuth T, Lupski JR. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res* 1991;19:6823-6831.
16. Gilson E, Clement JM, Brutlag D, Hofnung M. A family of dispersed repetitive extragenic palindromic DNA sequences in *E.coli*. *EMBO J* 1984;3:1417-1421.
17. Hulton CS, Higgins CF, Sharp PM. ERIC sequences: a novel family of repetitive elements in the genome of *Escherichia coli*, *Salmonella typhimurium* and other enterobacteria. *Mol Microbiol* 1991;5:8215-8834.
18. Koeuth T, Vesalovic J, Lupski JR. Differential subsequence conservation of interspersed repetitive *Streptococcus pneumoniae* BOX elements in diverse bacteria. *Genome Res* 1995;5:408-418.
19. Ramazanzadeh R, McNerney R. Variable number of tandem repeats (VNTR) and its application in bacterial epidemiology. *Pakistan J Biol Sci* 2007;10:2612-2621.
20. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *P Natl Acad Sci USA* 1977;74:5463-5467.
21. Ahmadian A, Ehn M, Hober S. Pyrosequencing: history, biochemistry and future. *Clinica Chimica Acta* 2006;363:83-94.
22. Jensen MA, Webster JA, Stratus N. Rapid identification of bacteria on the basis of polymerase chain reaction-amplified ribosomal DNA spacer polymorphism. *Appl Environ Microbiol* 1993;59:945-952.
23. [http://en.wikipedia.org/wiki/Housekeeping\\_gene](http://en.wikipedia.org/wiki/Housekeeping_gene) (Accessed 01/06/2011).









