http://www.pjbs.org

PJBS

ISSN 1028-8880

Pakistan Journal of Biological Sciences

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Pakistan Journal of Biological Sciences 10 (16): 2612-2621, 2007
ISSN 1028-8880

Variable Number of Tandem Repeats (VNTR) and its Application in Bacterial Epidemiology

[1]Rashid Ramazanzadeh and [2]Ruth McNerney
[1]Department of Microbiology, Faculty of Medicine,
Kurdistan University of Medical Science, Sanandaj, Iran
[2]London School of Hygiene and Tropical Medicine, London WC1E 7HT, United Kingdom

Abstract: Molecular epidemiology is the using of molecular techniques to study bacterial distribution in human populations. Recently molecular epidemiologist benefit from several techniques such as Variable Number Tandem Repeat (VNTR) typing method to typing bacterial strains. Variable Number Tandem Repeat (VNTR) typing is a tool for genotyping and provides data in a simple and numeric format based on the number of repetitive sequences. VNTR for first time identified in *M. tuberculosis* as Mycobacterial Interspersed Repeat Units (MIRUs). General terms of VNTR have now been reported in Bacillus anthracis, *Legionella pneumophila*, *Pseudomonas aeruginosa*, *Salmonella enterica* and *Escherichia coli* O157.

Key words: VNTR, repetitive DNA, MIRU, spoligotyping, molecular epidemiology

INTRODUCTION

Tuberculosis remains a major public health threat worldwide and the mortality rate is about 3 million people every year despite of availability of antituberculosis drugs for almost 50 years (Narayanan, 2004). Estimated that 1/3 of the human population is infected by causative agent *Mycobacterium tuberculosis*, (Mazars *et al*., 2001) and annually about 10 million new cases of active tuberculosis, is created by this agent, which indicate existence of huge reservoir for the disease (Malik and Godfrey-Faussett, 2005; Mazars *et al*., 2001; Supply *et al*., 2001). Since 1882 that *Mycobacterium tuberculosis* was described as the causative agents of tuberculosis by Koch (Malik and Godfrey-Faussett, 2005) other members of the *M. tuberculosis* complex (*Mycobacterium africanum*, *Mycobacterium bovis*, *Mycobacterium microti*) also cause disease in humans and animals. Although strains of the *M. tuberculosis* complex vary in host range, virulence and other phenotypes, they have highly conserved DNA sequences (Frothingham and Meeker-O'Connell, 1998).

Normally tuberculosis (TB) patients infected with a single *Mycobacterium tuberculosis* strain and infection with one strain could be able to confer immunity to additional *M. tuberculosis* infections (Shamputa *et al*., 2004) but the occurrence of mixed infections has now become generally accepted. Today recurrence (a disease that caused by original strain after treatment), reinfection (cases of infection by a second *M. tuberculosis* strain) and mixed infection (infection with more than one strains of *M. tuberculosis*) made a difficulty interpretation of Drug Susceptibility Testing (DST) results and the detection of epidemiological links (Shamputa *et al*., 2004). In addition development of disease state can be affected by several well-known host risk factors, which include malnutrition, alcoholism, homelessness and occupational lung disease, vitamin D deficiency, NRAMP1(natural resistance-associated macrophage protein one) polymorphisms, combination of demographic factors, population movements, HIV infection and multiple drug resistance strains (Malik and Godfrey-Faussett, 2005; (Kanduma *et al*., 2003; Mazars *et al*., 2001). The frequency of different clinical presentations of tuberculosis is also different between ethnic groups. The worldwide development of transport and migration contributes to globalize those threats (Mazars *et al*., 2001). Because of these factors, efficient disease control programs are needed by international epidemiological surveillance systems that able to monitor worldwide transmission of Tuberculosis (Supply *et al*., 2001).

In order to gain a better understanding of factors that influence TB transmission in a community and for evaluation of regional control programmes strain identification can be used as useful tool in epidemiological investigations. Strain identification is doing an essential trend by standardized molecular fingerprinting techniques that can be used for comparison

Corresponding Author: Rashid Ramazanzadeh, Department of Microbiology, Faculty of Medicine, Kurdistan University of Medical Science, Sanandaj, Iran

of strains between laboratories, regions, countries and continents (Kanduma *et al*., 2003). In addition strain identification can be useful for determine the outbreak strain and distinguishing it from epidemiologically unrelated strain. All isolate that involved in outbreaks of an infection would be expected to be clonal. Clonally related family of *M. tuberculosis* have been determined by molecular-typing methods which may be restricted to special area or disseminated worldwide (Supply *et al*., 2001). Beijing family is an example that identified in Beijing, China, in 1992 (Mokrousov *et al*., 2004) and then reported from most part of the world. Several studies indicate that this family strains associated with drug resistance and high adaptability to the host intracellular environment. This family strains often carry more than 20 copies of IS6110 in their genome, with a high degree of similarity in band pattern, which makes computer-assisted analysis particularly difficult and cumbersome (Kam *et al*., 2005) and genetically, these strains are closely related and thus difficult to differentiate by most of the currently used typing techniques (Mokrousov *et al*., 2004), therefore needed precious molecular technique to define this strains. With the advent of molecular techniques, TB investigators have new and powerful tools to further understand the transmission and phylogenetic properties of *M. tuberculosis*. This paper will review new molecular epidemiological tools specially VNTR and use of this technique in other bacteria and also the application of this technique to answer epidemiological, clinical and biological questions.

MOLECULAR EPIDEMIOLOGY

Molecular epidemiology is the study of distribution and determinants of disease occurrence in human populations using molecular techniques. Therefore is a blend of molecular biology and epidemiology (Narayanan, 2004). The applications of Molecular techniques are:

- To determine the risk factors for the transmission of the disease (e.g., TB transmission in a community).
- To distinguish between exogenous (recent infection) and endogenous (reactivation) disease to investigate outbreaks and to identify the source of laboratory cross contamination (Easterbrook *et al*., 2004; Bifani *et al*., 2000).
- To study transmission within a defined geographical setting.
- To demonstrate the occurrence of exogenous superinfection in immunocompetent and immunocompromised patients and exogenous re-infection after curative treatment.
- Can be useful in informing the control measures necessary to break the chain of transmission of Multi-Drug Resistance-TB (MDRTB) (Kanduma *et al*., 2003).
- To track global spread of pathogens e.g., TB.
- understand the virulence and resistance mechanisms of different strains.
- To improve the knowledge on transmission dynamics and dissemination pathways of infectious diseases.
- To develop strategies for the treatment and the prevention of the tuberculosis (Tazi *et al*., 2002).
- To study the spread of tubercle bacilli in mini epidemics, outbreaks.
- To track the geographic distribution and spread of clones of *M. tuberculosis* of public health importance (Narayanan, 2004; Malik and Godfrey-Faussett, 2005).
- To measure the impact of any public health intervention strategy for control of tuberculosis in the community (Easterbrook *et al*., 2004).

Molecular typing methods started from 1990 with restriction fragment length polymorphism (RFLP) analysis based on the insertion sequence IS6110 and then developed to PCR based methods. Recently microarray based methods are also available. These methods will enable individual strains of *M. tuberculosis* or clonal groups to be identified by specific phenotypic traits (Burgos and Pym, 2002).Validity of each DNA typing method determined by, typeability, reproducibility, stability and discriminatory power (Blackwood *et al*., 2004; Burgos and Pym, 2002). With considering that all current typing methods suffer from significant drawbacks but large-scale genotyping of *M. tuberculosis* is necessary for determining the worldwide transmission of TB. To achieve this goal, effective methods for accurate identification and monitoring of large numbers of *M. tuberculosis* strains are required (Mazars *et al*., 2001).

NON-DNA TYPING METHODS

Before development of molecular techniques, biochemical features, susceptibility to antibiotics and phage typing were the most used methods for differentiation of strains of the TB complex. Serological methods have been used for diagnosis but do not differentiate between infections with different strains of *M. tuberculosis*. However, phage typing is valuable epidemiological tool in typing *M. tuberculosis* strains from outbreaks and laboratory cross examination (McNerney, 1999). Disadvantages of these methods are cumbersome and low sensitivity and limited number of possible patterns (Kanduma *et al*., 2003).

Traditional methods are often based on phenotypic characteristics and include biotyping, serotyping, antibiogram analysis, bacteriocin typing, phage typing and plasmid typing. These traditional methods had been found to be insufficiently discriminatory, have poor reproducibility, suffer from lack of availability of specific reagents and are affected by physiological factors Alfizah, 2004).

GENOTYPING (DNA FINGERPRINTING TECHNIQUES)

Genome-based typing: Publication of complete genome of *M. tuberculosis* in 1998 made new opportunity to screening of the genome for genetics polymorphisms as a candidate that contributes to pathogenicity and epidemiology. Therefore, differentiation of strains based on genetic polymorphisms has been providing a new more direct approach of exploring strain variation (Malik and Godfrey-Faussett, 2005). Several molecular technique had been invented by using of genome including Restriction Fragment Length Polymorphism (RFLP) typing, which remains the gold standard, spacer oligonucleotide typing (spoligotyping) and mycobacterial interspersed units typing (MIRU typing). Although many other techniques have been used in laboratory practice (Malik and Godfrey-Faussett, 2005).

Pulsed field gel electrophoresis: Pulsed-Field Gel Electrophoresis (PFGE) typing was described by Schwartz and Cantor (1984). PFGE is highly effective in molecular epidemiological studies of bacterial isolates and allow the electrophoretic separation of larger molecules in agarose gels. This method is used to discriminating among isolates such as *Escherichia coli*, *Staphylococcus aureus* and many other species. Investigation of clonal relatedness of outbreaks isolates can be done with using this technique (Kanduma *et al*., 2003; Alfizah, 2004).The limit of resolution of PFGE depends on several factors, including:

- The uniformity of the two electric fields,
- The duration of the electric pulses,
- The ratio of the pulse times for each of the alternating electric fields,
- The angles of the two electric fields to the gel and
- The ratio of the strengths of the two electric fields (Maiden and Pollard, 2002).

The lack of standardized methods to generate, store and compare PFGE patterns has also limited the use of PFGE for population-based molecular epidemiologic studies of *M. tuberculosis* (Singh, 1999).

Restriction fragment length polymorphism: The first standardized protocol for Restriction Fragment Length Polymorphism (RFLP) typing of the Mycobacterium tuberculosis complex using IS6110 was published in 1993 (Thierry *et al*., 1993). This is usually demonstrated by digestion of the genomic DNA with specific restriction enzymes and analysis of the generated patterns after separation of the DNA fragments on agarose gel. Therefore strains of *M. tuberculosis* complex can be differentiating based on strain specific differences and frequencies of certain DNA sequences in chromosomal DNA. This kind of analysis is technically possible and no hybridization step with defined probes is needed. However, interpretation of the results is difficult because the large number of fragments generates a complex pattern and only a small number of different RFLP types are observed (Kanduma *et al*., 2003).

DNA microarrays: A DNA microarray is a solid substrate containing DNA or oligonucleotides that can be hybridize with complementary labeled nucleic acid molecules (Cummings and Relman, 2000). Application of DNA microarray technology is to detect global patterns of gene expression, identification of complex genetic diseases, mutation/polymorphism detection, drug discovery/ toxicology studies and tissue microarrays that used for the molecular profiling of tumor specimens (Dolan *et al*., 2001). The advantage of microarrays is sensitivity and detection the expression of many thousands of genes (20, 000-40 000) in one experiment. The limitation of this technique is that only can be used for fully sequenced isolates and needed software for analysis (Malik and Godfrey-Faussett, 2005).

Deletion-based typing: Deligotypng is a PCR based method same as spoligotyping. In this method the presence or absence of specific deletions by hybridisation of the test strain to a membrane bound with deletion-specific oligonucleotides. This method provides a new typing system for further insight to understanding the molecular epidemiology of *M. tuberculosis* (Malik and Godfrey-Faussett, 2005; Goguet de la Salmoniere *et al*., 2004).

Synonymous single-nucleotide polymorphisms: Bacteria have simple structure and prone to exchange the genetic contents gene deletion or by acquisition of mobile elements such as plasmids and bacteriophages. Two classes of substitutions, referred to as synonymous and nonsynonymous Single Nucleotide Polymorphisms (SNPs), can occur in genes that encode proteins. Nonsynonymous SNPs (nsSNPs) result in amino acid replacements and hence provide substrate for

evolutionary selection. Synonymous Single-nucleotide Polymorphisms (sSNPs) do not alter the structure of proteins and are therefore functionally neutral. SNPs are easy to detect and they provide useful targets for large-scale molecular population genetic studies examining evolutionary relationships among bacterial strains, especially in strongly clonal species. SNPs are used in phylogenetic studies of *M. tuberculosis* (Malik and Godfrey-Faussett, 2005; Gutacker *et al*., 2002; (Fleischmann *et al*., 2002).

Hybridization-based methods

RFLP- IS6110: IS6110 belongs to the IS3 family of insertion sequences and contain 1,355 base pairs. It has been described in 1989 for the first time. The range of copy number of this marker is 0-25. The positions of IS6110 highly variable between different isolates of *M. tuberculosis*. This variability is sufficient to generate RFLP and for it to be used in fingerprinting. Therefore the isolate can be differentiated based on the copy number and position of this sequence. IS6110-based typing is the most widely applied genotyping method in the molecular epidemiology of *M. tuberculosis* (Burgos and Pym, 2002). Though IS6110 RFLP typing is the Gold standard for typing strains of *M. tuberculosis*, it has several disadvantages. It is a slow, cumbersome, labour intensive (requiring a large culture biomass, technical expertise and time) and its discriminatory power for isolates with fewer than six IS6110 copies is low and, consequently, secondary typing using another independent genetic marker is often required (Narayanan, 2004), (Blackwood *et al*., 2004; Mazars *et al*., 2001). Also IS6110-RFLP profiles difficult to compare between intralaboratory or interlaboratory and needed specialized software's (Sola *et al*., 2001; Supply *et al*., 2001; Sun *et al*., 2004).

Polymorphic guanine-cytosine-rich repetitive sequence restriction fragment length polymorphism typing: Polymorphic guanine-cytosine-rich repetitive sequence restriction fragment length polymorphism typing (PGRS-RFLP) has been shown to have a discriminatory power close to that of IS6110 typing, even in isolates with low copy numbers of IS6110. Therefore it can be used as secondary typing system. However, the PGRS regions comprise many nonperfect repeats, making the RFLP patterns complex and sometimes difficult to interpret (Burgos and Pym, 2002).

PCR based-methods: PCR-based methods are easier to perform, require relatively smaller amounts of genomic DNA and even can be performed on non viable organisms or directly from clinical specimens relative to RFLP genotyping. Recently many PCR based typing assays have been developed (Burgos and Pym, 2002).

Ligation mediated PCR (LMPCR): Ligation mediated PCR is amplification of a flanking sequence located on the 5' side of IS6110. This method gives molecular patterns and technically simple and reproducible. The LM-PCR patterns were normalized and subjected to cluster analysis using the Bionumerics software (Niobe-Eyangoh *et al*., 2003, Bonora *et al*., 1999; Prod'hom *et al*., 1997). However this method is based on IS6110 element and hence not useful for typing the isolates with low copy numbers of IS6110 (Burgos and Pym, 2002).

Exact tandem repeats (ETR): Exact tandem repeats (ETRs) have also been used for PCR-based strain typing assays. ETRs are variable number of tandem repeats ranging from 53 to 79 bp in length, which vary between strains and between different species of the *M. tuberculosis* complex (Frothingham and Meeker-O'Connell, 1998).The discriminatory power of this method was not as good as that achieved using spoligotyping or IS6110-RFLP typing (Roring *et al*., 2002).

Amplified fragment length polymorphism (AFLP): Amplified Fragment Length Polymorphism (AFLP) analysis is a PCR-based technique involving the restriction of bacterial genomic DNA with two restriction enzymes, the ligation of adapters to the restriction fragments and selective amplification of sets of the restriction fragments with adapter-specific primers. The amplified fragments are then sized by gel electrophoresis. The addition of a fluorescent label to one of the PCR primers allows the amplified fragments to be detected with an automated DNA sequencer (Frothingham and Meeker-O'Connell, 1998). Recently complete genome sequence of *M. tuberculosis* available design genome-based FAFLP conditions to establish strain types for this pathogen and to determine genetic relationships between clinical isolates (Goulding *et al*., 2000).

Spoligotyping: Spoligotyping is a Polymerase Chain Reaction (PCR)-based method that interrogates a small DR sequence with 36 bp repeats interspersed with short unique, non repetitive sequences 35-71 bp in length. All these spacer nucleotides between the direct repeats can be amplified simultaneously using one set of primers. The presence or absence of spacers in a given biotinylated strain is determined by hybridization with a set of 43 oligonucleotides derived from spacer sequences of *M. tuberculosis* H37Rv. The presence or absence of spacers is then determined via Southern hybridisation.

Individual strains are distinguished by the number of spacers missing from the complete spacer set defined by sequencing this region from a large number of *M. tuberculosis* strains.

The DR locus is presently the only well-studied single locus in the genome of *M. tuberculosis* showing considerable strain-to-strain polymorphism. The nature of polymorphism has been used to genotypically classify clinical isolates by DR-RFLP to define epidemiological relationships. Although the overall discriminatory power of spoligotyping is lower than that of IS6110 typing, it has the specific advantage of higher discrimination of strains with low copy numbers of IS611016 (Narayanan, 2004). This method requires very small amounts of DNA, which can be obtained without bacterial culture. In addition, PCR can be carried out directly on clinical specimens, enabling simultaneous identification and strain typing of *M. tuberculosis* in sputum from patients (Frothingham and Meeker-O'Connell, 1998; Goulding *et al*., 2000). Another advantage of this secondary marker technique is that it is economical, easy to perform and a rapid means of typing the *M. tuberculosis* complex. These characteristics make it a candidate for use in resource poor situations. Since *M. tuberculosis* isolates with a different spoligotype invariably have distinguishable IS6110 profiles, spoligotyping could conceivably be used as an initial screening step before applying a secondary technique of greater discriminatory power (Burgos and Pym, 2002, Sola *et al*., 2001).

VNTR: Tandemly repeated sequences are dispersed by thousands of copies in virtually all higher eukaryote genomes. Loci with Short Sequence Repeats (SSRs) of 1-13 bp are generally referred to as microsatellites and those with 10-100 bp sequence repeats as minisatellites. Many of these loci show hypervariability in their repeat numbers in humans and in animals and are therefore also called Variable Number Tandem Repeat (VNTR) loci. Variable Number Tandem Repeat (VNTR) typing is a tool for genotyping and it provides data in a simple and nonambiguous format based on the number of repetitive sequences. However, many aspects of the mechanisms of the variability, its biological roles and its involvement in evolutionary processes still remain mysterious (Mazars *et al*., 2001; Supply *et al*., 2000). The tandem repeat nature of minisatellites has been studied extensively in eukaryotic genomes. The first bacterial species in which they were identified was *Mycobacterium tuberculosis* that is being described as Mycobacterial Interspersed Repeat Units (MIRUs). These loci have a characteristic structure consisting of a variable number of near-identical repeated DNA sequences arranged consecutively. A general term for these repeats is Variable-Number Tandem Repeats (VNTRs) and true VNTRs have now been reported in *Bacillus anthracis*, *Legionella pneumophila*, *Pseudomonas aeruginosa*, *Salmonella enterica* and *Escherichia coli O157*.

Mycobacterial Interspersed Repeat Units (MIRU): MIRUs are short (40-100 bp) DNA elements often found as tandem repeats and dispersed in intergenic regions in the genome of the *M. tuberculosis* complex. The strains vary in the number of repeats at different loci. Each typed strain is assigned a 12-digit number corresponding to the number of repeats at each MIRU loci, forming the basis of a coding system that facilitates interlaboratory comparisons. The technical difficulty of sizing the multiple small PCR fragments is overcome by combining multiplex PCR with a fluorescence-based DNA analyzer. MIRU VNTR profiling is fast, appropriate for strains regardless of their IS6110 RFLP copy number and permits rapid comparison of global strains using a binary data classification system. This 12-loci-based approach allows for direct and reliable comparison of results between laboratories and thus for construction and analysis of global databases via the Internet for large-scale epidemiological and population genetic studies (Blackwood *et al*., 2004; Overduin *et al*., 2004; Sun *et al*., 2004; Mazars *et al*., 2001; Supply *et al*., 2000; Supply *et al*., 2001). MIRU typing also showed its usefulness in studying the population structure of Beijing family of *M. tuberculosis* (Kam *et al*., 2005).This new method was shown with different strain samples to possess a higher discriminatory power than that of spoligotyping and only slightly below that of IS6110-RFLP typing. The apparent advantage of the MIRU approach (compared to the IS6110 typing) is its portability due to easy digitalization of the generated profiles and hence easy interlaboratory exchange, as well as easy creation and maintenance of the databases (Mokrousov *et al*., 2004).

Multiple-locus variable-number tandem repeat analysis (MLVA) in Enterococci: *Enterococci* are gram-positive, spherical bacteria that colonize in groups or chains. They are naturally found as part of the digestive tract flora in many organisms, including humans. They are robust microbes able to tolerate relatively high salt and acid concentrations. Currently, *Enterococcus* infections account for 12% of all nosocomial infections, second only to *E. coli*. An *Enterococcus* infection can cause complicated abdominal infections, skin (Noskin *et al*., 1995) and skin structure infections, urinary tract infections and infections of the blood stream. These infections can

be difficult to treat, particularly in cases where the strain involved has developed resistance to several antibiotics. There are two species of *Enterococci*, *E. faecalis* and *E. faecium* is known to have a resistance to several types of antibiotics including quinolones and aminoglycosides (Abele-Horn *et al*., 2006; Zubaidah *et al*., 2006; Nicas *et al*., 1989). Multiple-locus variable-number tandem repeat (VNTR) analysis (MLVA) is a novel typing method to assess the genetic relatedness of *E. faecium* isolates with using of six VNTR loci. Furthermore, the discriminatory power of MLVA is comparable to Multilocus Sequence Typing (MLST). MLVA is a highly reproducible and portable typing method; in contrast to MLST, it is fast, relatively cheap and easy to perform. Furthermore, it has the abilities of MLST to recognize genetically related and potential epidemic isolates. Submission of MLVA profiles is possible via a Web-based database (http://www.mlva.umcutrecht.nl) for international comparison (Top *et al*., 2004). In another study Multilocus Variable-number Tandem-repeat Analysis (MLVA) for seven genomic loci was developed for *Enterococcus faecalis*. This method compared with Pulsed-field Gel Electrophoresis (PFGE).The data indicated that both typing schemes were highly concordant (90.4%) and MLVA is an excellent alternative to pulsed-field gel electrophoresis (PFGE) (Titze-de-Almeida *et al*., 2004).

VNTR application in paratuberculosis: Paratuberculosis, also named Johne's disease in ruminants, is characterized by a chronic inflammation of the ileum and is caused by *Mycobacterium avium* subsp. *paratuberculosis*. In humans, the symptoms of Crohn's disease partly resemble those of Johne's disease in ruminants. However, the etiology of Crohn's disease is much more complex and appears to be multifactorial. Because of the high prevalence of these bacteria in (products of) ruminants and their remarkable thermostability, concern has been raised about the possible role of these bacteria in the pathogenesis of Crohn's disease. In an attempt to develop a molecular typing method to facilitate meaningful comparative DNA fingerprinting of *M. avium* subsp. *paratuberculosis* isolates from the human and animal reservoirs, multilocus variable-number tandem-repeat analysis (MLVA) was explored and compared to IS900 restriction fragment length polymorphism (RFLP) typing. MLVA typing subdivided the most predominant RFLP type, R01, into six subtypes and thus provides a promising molecular subtyping approach to study the diversity of *M. avium* subsp. *Paratuberculosis* (Overduin *et al*., 2004).

VNTR application in Haemophilus influenza type b: *Haemophilus influenza* is a Gram-negative bacterium. Serious infection is usually caused by strains carrying a polysaccharide capsule. Of the six capsular types, type b (Hib) causes almost all systemic infections including; meningitis epiglottitis, osteomyelitis, septic arthritis and septicaemia. Up to 15% of children in non-immunized populations may harbour Hib in their nasopharynx. *H. influenza* type b (Hib) is estimated to cause at least 3 million cases of serious disease and 400 000-700, 000 deaths each year in young children (Sheff, 2006; Sinsalo *et al*., 2002). PCR-mediated amplification of Variable Number of Tandem Repeat Regions (VNTRs), have been used to detect the epidemic strain. The results show that VNTRs comprising repeat units that are 3, 5, or 6 nucleotides in length provided stable genetic markers that can be used for molecular typing of *H. influenza* type b. VNTRs built from tetranucleotide units, however, appear to be hyper variable and not suited for epidemiological studies (van Belkum *et al*., 1997).

Multiplelocus VNTR analysis (MLVA) application in *Salmonella enterica* subsp. *enterica*: Multiplelocus VNTR analysis (MLVA) based on 10 VNTRs have been used to genotype the *Salmonella enterica* subsp. *enterica* strains. Cluster analysis not only identified three genetically distinct groups consistent with the present serovar classification of salmonellae (serovars *Typhi*, *Paratyphi A* and *Typhimurium*) but also discriminated 25 subtypes within serovar Typhi isolates. Therefore, MLVA has potential to use in outbreaks investigations of typhoid fever (Ramisse *et al*., 2004).

MLVA application in *Francisella tularensis*: *Francisella tularensis* is a small Gram-negative aerobic bacillus with two main serotypes: Jellison Type A and Type B. Type A is the more virulent form *F. tularensis*, the etiological agent of tularemia, is found throughout the Northern hemisphere. After analyzing the *F. tularensis* genomic sequence for potential Variable-number Tandem Repeats (VNTRs), a multilocus VNTR analysis (MLVA) typing system based on six VNTR was developed for this pathogen. PCR assays revealed diversity at these loci with total allele numbers ranging from 2 to 20. Cluster analysis identified two genetically distinct groups consistent with the current biovar classification system of *F. tularensis*. These findings suggest that these VNTR markers are useful for identifying *F. tularensis* isolates at this taxonomic level (Farlow *et al*., 2001; Supply *et al*., 2000).

VNTR application in *Bacillus anthracis*: Bacillus represents a genus of Gram-positive bacteria which are ubiquitous in nature (soil, water and airborne dust). *B. anthracis* is the bacterium which causes anthrax in cows, sheep and sometimes humans. Anthrax is transmitted to humans via direct contact with animal products or inhalation of endospores. Under the microscope, *B. anthracis* cells appear to have square ends and seem to be attached by a joint to other cells. The spores are best observed when the bacterium is cultured on artificial media. *Bacillus anthracis* is one of the most genetically homogeneous pathogens described, making strain discrimination particularly difficult. Multiple-locus VNTR analysis (MLVA) uses the combined power of multiple alleles at several marker loci. Recently, fluorescently labeled PCR primers are used to produce PCR amplification products from eight VNTR regions in the *B. anthracis* genome. These are detected and their sizes are determined using an ABI377 automated DNA sequencer. Five of these eight loci were discovered by sequence characterization of molecular markers (*vrrC*1, *vrrC*2, *vrrB*1, *vrrB*2 and CG3), two were discovered by searching complete plasmid nucleotide sequences (pXO1-aat and pXO2-at) and one was known previously (*vrrA*) (Keim *et al*., 2000). *Bacillus anthracis* contains 30 comparable structures in which the unit is repeated at least 10 times. Half of these tandems repeat show polymorphism among the strains tested (Le Fleche *et al*., 2001).

VNTR application in *Yersinis pestis*: The Gram-negative bacterium *Yersinia pestis* is the causative agent of the systemic invasive infectious disease classically referred to as plague and has been responsible for three human pandemics: the Justinian plague (sixth to eighth centuries), the Black Death (fourteenth to nineteenth centuries) and modern plague (nineteenth century to the present day). The recent identification of strains resistant to multiple drugs and the potential use of *Y. pestis* as an agent of biological warfare mean that plague still poses a threat to human health. *Yersinia pestis* contains 64 such minisatellites in which the unit is repeated at least 7 times. An additional collection of 12 loci with at least 6 units and a high internal conservation were also evaluated. Forty-nine are polymorphic among five *Yersinia* strains (twenty-five among three *Y. pestis* strains). Analysis of the currently available bacterial genome sequences classifies *Yersinia pestis* as having an average density of tandem repeat arrays longer than 100 bp when compared to the other bacterial genomes analyzed to date. In these cases, testing a fraction of these sequences for polymorphism was sufficient to quickly develop a set of more than fifteen informative markers, some of which show a very high degree of polymorphism (Le Fleche *et al*., 2001).

Staphylococcal interspersed repeat units (SIRUs): Staphylococci are a common type of bacteria that live on the skin and mucous membranes (e.g., in nose) of humans. *Staphylococcus aureus* is the most important of these bacteria in human diseases. Other staphylococci, including *S. epidermidis*, are considered commensals, or normal inhabitants of the skin surface. Variable-number Tandem Repeats (VNTRs) have been shown to be a powerful tool in the determination of evolutionary relationships and population genetics of bacteria. The sequencing of a number of *Staphylococcus aureus* genomes has allowed the identification of novel VNTR sequences in *S. aureus*, which are similar to those used in the study of the evolution of *Mycobacterium tuberculosis*. Seven VNTRs, termed staphylococcal interspersed repeat units (SIRUs), distributed around the genome that occurring in both unique and multiple sites and varying in length from 48 to 159 bp. Variations in copy numbers were observed in all loci. Which those VNTRs may be a more appropriate evolutionary marker for studying transmission events and the geographical spread of *S. aureus* strains (Hardy *et al*., 2004).

VNTR application in E. coli O157:H7: *Escherichia coli* O157:H7 is a major cause of food-borne disease. Outbreak detection involves traditional epidemiological methods and routine molecular subtyping by Pulsed-field Gel Electrophoresis (PFGE). PFGE is labor-intensive and the results are difficult to analyze and not easily transferable between laboratories. Multilocus Variable-number Tandem Repeat (VNTR) Analysis (MLVA) is a fast, portable method that analyzes multiple VNTR loci, which are areas of the bacterial genome that evolve quickly. MLVA was performed by comparing the number of tandem repeats at seven loci. MLVA correctly identified the isolates from all outbreaks. MLVA differentiated strains with unique PFGE types. Additionally, MLVA discriminated strains within PFGE-defined clusters that were not known to be part of an outbreak. In addition to being a simple and validated method for *E. coli* O157:H7 outbreak detection, MLVA appears to have a sensitivity equal to that of PFGE and a specificity superior to that of PFGE (Noller *et al*., 2003).

VNTR application in *L. pneumophila*: *Legionella pneumophila* is the agent of Legionnaires' disease and Pontiac fever. This bacterium is present in aquatic environments, where it replicates within protozoan hosts. The bacteria infect alveolar macrophages following

inhalation of contaminated aerosols. The organism is responsible for a large number of cases of nosocomial pneumonia in immunocompromised patients. The family *Legionellaceae* contains more than 40 species that have been isolated from either clinical or environmental sources.Three subspecies of the species *L. pneumophila* have been described to be the most frequently associated with disease: *L. pneumophila* subsp. *pneumophila*, *L. pneumophila* subsp. *fraseri* and *L. pneumophila* subsp. *pascullei*. Different molecular techniques have been developed to characterize and analyze the different strains. Macrorestriction analysis and PCR-based methods such as Amplified Fragment Length Polymorphism (AFLP) analysis and arbitrarily primed PCR have been used to genotype *L. pneumophila* isolates. The availability of genome sequences allows the search for new markers. Based upon the genome sequence of the Philadelphia-1 strain (serogroup 1), 25 minisatellites were selected and their polymorphisms were analyzed by PCR. Most of these markers did not amplify the DNA of other *Legionella* species or other bacteria used as controls. A polymorphism was observed for seven markers among the *L. pneumophila* strains tested. VNTR typing is as informative as pulsed-field gel electrophoresis for comparison of strains. This method allowed the rapid identification of the *L. pneumophila* colonies and gave a first hint as to the presence of several strains in a single sample (Pourcel *et al*., 2003).

CONCLUSION

Molecular epidemiology made new vision in tuberculosis research, prevention and control. ME permit differentiation between strains by region, population and other risk factors and evaluation the prevalence of endemic strains. Molecular methods based on several markers are necessary for better strain definition. The gold standard method is Restriction Fragment Length Polymorphism (RFLP) analysis based on the insertion sequence IS6110 that produce unique pattern. The other genetic method is spoligotyping based on Direct Repeat (DR) polymorphism. Discriminatory power of this method is lower than IS6110-RFLP but spoligotyping is more suitable for investigating the biogeographic distribution of families of *M. tuberculosis* complex strains. Variable-number tandem repeats, named mycobacterial interspersed repetitive units (MIRU-VNTR), are another type of variable element in the *M. tuberculosis* complex genome showing extensive polymorphism with a discrimination power close to that of IS*6110*-RFLP. Because of their stability, they can be used for a clear definition of families of tubercle bacilli as well. VNTR can be used to define clonal expansion other bacteria such as *E. coli*, *S. aureus*, *L. pneumophila*, *Y. pestis*, *B. anthracis*, *F. tularensis*, *S. enterica* subsp. *Enterica* and *Enterococci*.

Several molecular epidemiology methods indicated that population structure of bacterial species appears to be strictly, or at least predominantly, clonal which disseminated locally or globally. Molecular epidemiology is a powerful approach to monitoring infectious diseases. All of mentioned methods have some disadvantages the make difficulties in correct evaluation of infectious disease. In this paper the overall aspect of VNTR as a new approach in molecular methods has been mentioned. Therefore more investigation is required to develop the robust, rapid, cheap and more valuable molecular tools for epidemiological research.

REFERENCES

Abele-Horn, M., U. Vogel, I. Klare, C. Konstabel, R. Trabold, R. Kurihara, W. Witte, W. Kreth, P.G. Schlegel and H. Claus, 2006. Molecular epidemiology of hospital-acquired vancomycin- resistant enterococci. J. Clin. Microbiol., 44: 4009-4013.

Alfizah, H.R.W., 2004. Using pulsed-field gel electrophoresis in the molecular investigation of an outbreak of *Serratia marcescens* infection in an intensive care unit. Singapore Med. J., 245: 214.

Bifani, P., S. Moghazeh, B. Shopsin, J. Driscoll, A. Ravikovitch and B.N. Kreiswirth, 2000. Molecular characterization of *Mycobacterium tuberculosis* H37Rv/Ra variants: Distinguishing the mycobacterial laboratory strain. J. Clin. Microbiol., 38: 3200-3204.

Blackwood, K.S., J.N. Wolfe and A.M. Kabani, 2004. Application of mycobacterial interspersed repetitive unit typing to *Manitoba tuberculosis* cases: Can restriction fragment length polymorphism be forgotten? J. Clin. Microbiol., 42: 5001-5006.

Bonora, S., M.C. Gutierrez, G. Di Perri, F. Brunello, B. Allegranzi, M. Ligozzi, R. Fontana, E. Concia and V. Vincent, 1999. Comparative evaluation of ligation-mediated PCR and spoligotyping as screening methods for genotyping of *Mycobacterium tuberculosis* strains. J. Clin. Microbiol., 37: 3118-3123.

Burgos, M.V. and A.S. Pym, 2002. Molecular epidemiology of tuberculosis. Eur. Respir. J. Suppl., 36: 54s-65s.

Cummings, C.A. and D.A. Relman, 2000. Using DNA microarrays to study host-microbe interactions. Emerg. Infect. Dis., 6: 513-525.

Dolan, P.L., Y. Wu, L.K. Ista, R.L. Metzenberg, M.A. Nelson and G.P. Lopez, 2001. Robust and efficient synthetic method for forming DNA microarrays. Nucl. Acids Res., 29: E1077.

Easterbrook, P.J., A. Gibson, S. Murad, D. Lamprecht, N. Ives, A. Ferguson, O. Lowe, P. Mason, A. Ndudzo, A. Taziwa, R. Makombe, L. Mbengeranwa, C. Sola, N. Rastogi and F. Drobniewski, 2004. High rates of clustering of strains causing tuberculosis in Harare, Zimbabwe: A molecular epidemiological study. J. Clin. Microbiol., 42: 4536-4544.

Farlow, J., K.L. Smith, J. Wong, M. Abrams, M. Lytle and P. Keim, 2001. Francisella tularensis strain typing using multiple-locus, variable-number tandem repeat analysis. J. Clin. Microbiol., 39: 3186-3192.

Fleche, P., Y. Hauck, L. Onteniente, A. Prieur, F. Denoeud, V. Ramisse, P. Sylvestre, G. Benson, F. Ramisse and G. Vergnaud, 2001. A tandem repeats database for bacterial genomes: Application to the genotyping of Yersinia pestis and *Bacillus anthracis*. BMC Microbiology, 1: 2.

Fleischmann, R.D., D. Alland, J.A. Eisen, L. Carpenter, O. White, J. Peterson, R. DeBoy, R. Dodson, M. Gwinn, D. Haft, E. Hickey, J.F. Kolonay, W.C. Nelson, L.A. Umayam, M. Ermolaeva, S.L. Salzberg, A. Delcher, T. Utterback, J. Weidman, H. Khouri, J. Gill, A. Mikula, W. Bishai, W.R. Jacobs Jr, Jr., J.C. Venter and C.M. Fraser, 2002. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. J. Bacteriol., 184: 5479-5490.

Frothingham, R. and W.A. Meeker-O'Connell, 1998. Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats. Microbiol., 144: 1189-1196.

Goguet de la Salmoniere, Y.O., C.C. Kim, A.G. Tsolaki, A.S. Pym, M.S. Siegrist and P.M. Small, 2004. High-throughput method for detecting genomic-deletion polymorphisms. J. Clin. Microbiol., 42: 2913-2918.

Goulding, J.N., J. Stanley, N. Saunders and C. Arnold, 2000. Genome-sequence-based fluorescent amplified-fragment length polymorphism analysis of *Mycobacterium tuberculosis*. J. Clin. Microbiology, 38: 1121-1126.

Gutacker, M.M., J.C. Smoot, C.A. Migliaccio, S.M. Ricklefs, S. Hua, D.V. Cousins, E.A. Graviss, E. Shashkina, B.N. Kreiswirth and J.M. Musser, 2002. Genome-wide analysis of synonymous single nucleotide polymorphisms in *Mycobacterium tuberculosis* complex organisms: Resolution of genetic relationships among closely related microbial strains. Genetics, 162: 1533-1543.

Hardy, K.J., D.W. Ussery, B.A. Oppenheim and P.M. Hawkey, 2004. Distribution and characterization of Staphylococcal Interspersed Repeat Units (SIRUs) and potential use for strain differentiation. Microbiol., 150: 4045-4052.

Kam, K.M., C.W. Yip, L.W. Tse, K.L. Wong, T.K. Lam, K. Kremer, B.K. Au and D. van Soolingen, 2005. Utility of mycobacterial interspersed repetitive unit typing for differentiating multidrug-resistant *Mycobacterium tuberculosis* isolates of the Beijing family. J. Clin. Microbiol., 43: 306-313.

Kanduma, E., T.D. McHugh and S.H. Gillespie, 2003. Molecular methods for *Mycobacterium tuberculosis* strain typing: A users guide. J. Applied Microbiol., 94: 781-791.

Keim, P., L.B. Price, A.M. Klevytska, K.L. Smith, J.M. Schupp, R. Okinaka, P.J. Jackson and M.E. Hugh-Jones, 2000. Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. J. Bacteriol., 182: 2928-2936.Le

Maiden, M.C.J. and A.J.M. Pollard, 2002. Methods in Molecular Medicine, Meningococcal Disease: Methods and Protocols. Humana Press Inc. Totowa, NJ.

Malik, A.N. and P. Godfrey-Faussett, 2005. Effects of genetic variability of *Mycobacterium tuberculosis* strains on the presentation of disease. Lancet Infect. Dis., 5: 174-183.

Mazars, E., S. Lesjean, A.L. Banuls, M. Gilbert, V. Vincent, B. Gicquel, M. Tibayrenc, C. Locht and P. Supply, 2001. High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology. Proc. Natl. Acad. Sci .USA, 98: 1901-1916.

McNerney, R., 1999. TB: The return of the phage. A review of fifty years of mycobacteriophage research. Int. J. Tuberculosis Lung. Dis., 3: 179-184.

Mokrousov, I., O. Narvskaya, E. Limeschenko, A. Vyazovaya, T. Otten and B. Vyshnevskiy, 2004. Analysis of the allelic diversity of the mycobacterial interspersed repetitive units in *Mycobacterium tuberculosis* strains of the Beijing family: Practical implications and evolutionary considerations. J. Clin. Microbiol., 42: 2438-2444.

Narayanan, S., 2004. Molecular epidemiology of tuberculosis. Ind. J. Med. Res., 120: 233-247.

Nicas, T.I., C.Y. Wu, J.N. Hobbs, Jr., D.A. Preston and N.E. Allen, 1989. Characterization of vancomycin resistance in *Enterococcus faecium* and *Enterococcus faecalis*. Antimicrob Agents Chemother., 33: 1121-1124.

Niobe-Eyangoh, S.N., C. Kuaban, P. Sorlin, P. Cunin, J. Thonnon, C. Sola, N. Rastogi, V. Vincent and M.C. Gutierrez, 2003. Genetic biodiversity of *Mycobacterium tuberculosis* complex strains from patients with pulmonary tuberculosis in Cameroon. J. Clin. Microbiol., 41: 2547-2553.

Noskin, G.A., L.R. Peterson and J.R. Warren, 1995. *Enterococcusfaecium* and *Enterococcusfaecalis* bacteremia: Acquisition and outcome. Clin. Infect. Dis., 20: 296-301.

Noller, A.C., M.C. McEllistrem, A.G. Pacheco, D.J. Boxrud and L.H. Harrison, 2003. Multilocus variable-number tandem repeat analysis distinguishes outbreak and sporadic *Escherichia coli* O157:H7 isolates. J. Clin. Microbiol., 41: 5389-5397.

Overduin, P., L. Schouls, P. Roholl, A. van der Zanden, N. Mahmmod, A. Herrewegh and D. van Soolingen, 2004. Use of multilocus variable-number tandem-repeat analysis for typing *Mycobacterium avium* subsp. paratuberculosis. J. Clin. Microbiol., 42: 5022-5028.

Pourcel, C., Y. Vidgop, F. Ramisse, G. Vergnaud and C. Tram, 2003. Characterization of a tandem repeat polymorphism in *Legionella pneumophila* and its use for genotyping. J. Clin. Microbiol., 41: 1819-1826.

Prod'hom, G., C. Guilhot, M.C. Gutierrez, A. Varnerot, B. Gicquel and V. Vincent, 1997. Rapid discrimination of *Mycobacterium tuberculosis* complex strains by ligation-mediated PCR fingerprint analysis. J. Clin. Microbiol., 35: 3331-3334.

Ramisse, V., P. Houssu, E. Hernandez, F. Denoeud, V. Hilaire, O. Lisanti, F. Ramisse, J.D. Cavallo and G. Vergnaud, 2004. Variable number of tandem repeats in *Salmonella enterica* subsp. Enterica for typing purposes. J. Clin. Microbiol., 42: 5722-5730.

Roring, S., A. Scott, D. Brittain, I. Walker, G. Hewinson, S. Neill and R. Skuce, 2002. Development of variable-number tandem repeat typing of Mycobacterium bovis: Comparison of results with those obtained by using existing exact tandem repeats and spoligotyping. J. Clin. Microbiol., 40: 2126-2133.

Schwartz, D.C. and C.R. Cantor, 1984. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell, 37: 67-75.

Singh SP, L.C., M. Farid-Moyer and P.M. Small, 1999. Use of pulsed-field gel electrophoresis for molecular epidemiologic and population genetic studies of *Mycobacterium tuberculosis*. J. Clin. Microbiol., 37: 1927-1931.

Sinsalo, M., J. Aittoniemi, H. Kayhty and J. Vilpo, 2002. Haemophilus influenzae type b (Hib) antibody concentrations and vaccination responses in patients with chronic lymphocytic leukaemia: Predicting factors for response. Leuk Lymphoma, 43: 1967-1969.

Shamputa, I.C., L. R igouts, L.A. Eyongeta, N.A. El Aila, A. van Deun, A.H. Salim, E. Willery, C. Locht, P. Supply and F. Portaels, 2004. Genotypic and phenotypic heterogeneity among *Mycobacterium tuberculosis* isolates from pulmonary tuberculosis patients. J. Clin. Microbiol., 42: 5528-5536.

Sheff, B., 2006. Haemophilus influenzae Type B (Hib). Nursing, 36: 31.

Sola, C., S. Ferdinand, C. Mammina, A. Nastasi and N. Rastogi, 2001. Genetic diversity of *Mycobacterium tuberculosis* in Sicily based on spoligotyping and variable number of tandem DNA repeats and comparison with a spoligotyping database for population-based analysis. J. Clin. Microbiol., 39: 1559-1565.

Sun, Y.J., A.S. Lee, S.T. Ng, S. Ravindran, K. Kremer, R. Bellamy, S.Y. Wong, D. van Soolingen, P. Supply and N.I. Paton, 2004. Characterization of ancestral *Mycobacterium tuberculosis* by multiple genetic markers and proposal of genotyping strategy. J. Clin. Microbiol., 42: 5058-5064.

Supply, P., E. Mazars, S. Lesjean, V. Vincent, B. Gicquel and C. Locht, 2000. Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. Mol. Microbiol., 36: 762-771.

Supply, P., S. Lesjean, E. Savine, K. Kremer, D. van Soolingen and C. Locht, 2001. Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. J. Clin. Microbiol., 39: 3563-3571.

Tazi, L., B. Kreiswirth, C. Carriere and M. Tibayrenc, 2002. Molecular epidemiology of *Mycobacterium tuberculosis* and its relevance to the surveillance and control of TB: An e-debate. Infect. Genet. Evol., 2: 153-158.

Titze-de-Almeida, R., R.J. Willems, J. Top, I.P. Rodrigues, R.F. Ferreira, H. Boelens, M.C. Brandileone, R.C. Zanella, M.S. Felipe and A. van Belkum, 2004. Multilocus variable-number tandem-repeat polymorphism among Brazilian *Enterococcus faecalis* strains. J. Clin. Microbiol., 42: 4879-4881.

Thierry, D., P. Matsiota-Bernard, E. Pitsouni, C. Costopoulos and J.L. Guesdon, 1993. Use of the insertion element IS6110 for DNA fingerprinting of *Mycobacterium tuberculosis* isolates presenting various profiles of drug susceptibility. FEMS Immunol. Med. Microbiol., 6: 287-297.

Top, J., L.M. Schouls, M.J. Bonten and R.J. Willems, 2004. Multiple-locus variable-number tandem repeat analysis, a novel typing scheme to study the genetic relatedness and epidemiology of *Enterococcus faecium* isolates. J. Clin. Microbiol., 42: 4503-4511.

Van Belkum, A., W.J. Melchers, C. Ijsseldijk, L. Nohlmans, H. Verbrugh and J.F. Meis, 1997. Outbreak of amoxicillin-resistant Haemophilus influenzae type b: Variable number of tandem repeats as novel molecular markers. J. Clin. Microbiol., 35: 1517-1520.

Zubaidah, A.W., A. Ariza and S. Azmi, 2006. Hospital-acquired vancomycin-resistant enterococci: Now appearing in Kuala Lumpur Hospital. Med. J. Malaysia, 61: 487-489.