21 A genome-wide search for asthma susceptibility loci in ethnically diverse populations. The Collaborative Study on the Genetics of Asthma (CSGA). *Nat Genet* 1997; **15**: 389–392.
22 Li-Weber M, Krammer PH. Regulation of IL4 gene expression by T cells and therapeutic perspectives. *Nat Rev Immunol* 2003; **3**: 534–543.
23 Lee GR, Fields PE, Griffin TJ, Flavell RA. Regulation of the Th2 cytokine locus by a locus control region. *Immunity* 2003; **19**: 145–153.
24 Onnie C, Fisher SA, King K, Mirza M, Roberts R, Forbes A *et al*. Sequence variation, linkage disequilibrium and association with Crohn's disease on chromosome 5q31. *Genes Immun* 2006; **7**: 359–365.
25 Seah GT, Scott M, Rock GAW. Type 2 cytokine gene activation and its relationship to extend of disease in patients with tuberculosis. *J Infect Dis* 2000; **181**: 385–389.
26 Harris J, De Haro SA, Master SS, Keane J, Roberts EA, Delgado M *et al*. T helper 2 cytokines inhibit autophagic control of intracellular *Mycobacterium tuberculosis*. *Immunity* 2007; **27**: 505–517.
27 van Soolingen D, Qian L, de Haas PE, Douglas JT, Traore H, Portaels F *et al*. Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of East Asia. *J Clin Microbiol* 1995; **33**: 3234–3238.
28 Manca C, Reed MB, Freeman S, Mathema B, Kreiswirth B, Barry III CE *et al*. Differential monocyte activation underlies strain-specific *Mycobacterium tuberculosis* pathogenesis. *Infect Immun* 2004; **72**: 5511–5514.
29 Nuchnoi P, Ohashi J, Naka I, Nacapunchai D, Tokunaga K, Nishida N *et al*. Linkage disequilibrium structure of the 5q31-33 region in a Thai population. *J Hum Genet* 2008; **53**: 850–856.
30 Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD *et al*. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008; **40**: 955–962.
31 Peltekova VD, Wintle RF, Rubin LA, Amos CI, Huang Q, Gu X *et al*. Functional variants of OCTN cation transporter genes are associated with Crohn disease. *Nat Genet* 2004; **36**: 471–475.
32 van Crevel R, Karyadi E, Preyers F, Leenders M, Kullberg BJ, Nelwan RH *et al*. Increased production of interleukin 4 by CD4+ and CD8+ T cells from patients with tuberculosis is related to the presence of pulmonary cavities. *J Infect Dis* 2000; **181**: 1194–1197.
33 Vale RD. The molecular motor toolbox for intracellular transport. *Cell* 2003; **112**: 467–480.
34 Little KY, McLaughlin DP, Zhang L, Livermore CS, Dalack GW, McFinton PR *et al*. Cocaine, ethanol, and genotype effects on human midbrain serotonin transporter binding sites and mRNA levels. *Am J Psychiatr* 1998; **155**: 207–213.
35 Nishida N, Tanabe T, Takasu M, Suyama A, Tokunaga K. Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. *Anal Biochem* 2007; **364**: 78–85.
36 Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.
37 O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998; **63**: 259–266.
38 Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. *Bioinformatics* 2001; **17**: 742–743.
39 Lange C, DeMeo D, Silverman EK, Weiss ST, Laird NM. PBAT: tools for family-based association studies. *Am J Hum Genet* 2004; **74**: 367–369.
40 Luna A, Nicodemus KK. snp.plotter: an R-based SNP/haplotype association and linkage disequilibrium plotting package. *Bioinformatics* 2007; **23**: 774–776.
41 Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 2004; **74**: 765–769.
42 Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B *et al*. The structure of haplotype blocks in the human genome. *Science* 2002; **296**: 2225–2229.
43 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
44 Lee PH, Shatkay H. F-SNP: computationally predicted functional SNPs for disease association studies. *Nucleic Acids Res* 2008; **36**: D820–D824.
45 Dudbridge F. Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum Hered* 2008; **66**: 87–98.

Supplementary Information accompanies the paper on Genes and Immunity website (http://www.nature.com/gene)

Infection, Genetics and Evolution 10 (2010) 1066–1074

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Resolving lineage assignation on *Mycobacterium tuberculosis* clinical isolates classified by spoligotyping with a new high-throughput 3R SNPs based method

Edgar Abadia [a], Jian Zhang [a], Tiago dos Vultos [b], Viviana Ritacco [c], Kristin Kremer [d], Elif Aktas [e], Tomoshige Matsumoto [f], Guislaine Refregier [a], Dick van Soolingen [d], Brigitte Gicquel [b], Christophe Sola [a,b,*]

[a] Institute of Genetics and Microbiology, UMR8621, IGEPE Team, Universud, CNRS Université Paris-Sud 11, France
[b] Unité de Génétique Mycobactérienne, Institut Pasteur, Paris, France
[c] Instituto Nacional de Enfermedades Infecciosas ANLIS Carlos Malbrán, Vélez Sarsfield 563, 1281 Buenos Aires, Argentina
[d] National Institute for Public Health and the Environment, Bilthoven, The Netherlands
[e] Department of Clinical Microbiology, Zonguldak Karaelmas University Hospital, Zonguldak, Turkey
[f] Osaka Prefectural Hospital Organization Osaka Prefectural Medical Center for Respiratory and Allergic Diseases, Japan

ARTICLE INFO

Article history:
Received 19 May 2010
Received in revised form 1 July 2010
Accepted 2 July 2010
Available online 17 July 2010

Keywords:
Mycobacterium tuberculosis complex
CRISPR
Spoligotyping
Population studies
Phylogeny
Evolution
SNPs
Multiplex
High-throughput

ABSTRACT

We developed a new multiplexed-PCR assay to accurately classify *Mycobacterium tuberculosis* complex (MTC) isolates at the sublineage level by single nucleotide polymorphisms (SNPs). This method relies on 7 SNPs located in different genes of the MTC strains (*recC*, *recO*, *recR*, *ligB*, *ligC*, *alkA*, and *mgtC*). Most of these genes are involved in replication, repair and recombination (3R) functions of *M. tuberculosis* strains, four of the mutations are synonymous, and thus neutral. Genes were chosen as a first empirical approach to assess the congruence between spoligotyping-based phylogeographical classification and SNP typing.

This scheme efficiently classifies most of MTC phylogeographical groups: (1) confirming and identifying new sublineage-specific SNPs, (2) unraveling phylogenetical relationships between spoligotyping-defined MTC sublineages, (3) appropriately assigning sublineages to some spoligotypes and reassigning sublineages to other mis-labeled spoligotype signatures. This study opens the way to a more meaningful taxonomic, evolutionary and epidemiological classification. It also allows evaluation of spoligotype-signature significance towards a more comprehensive understanding of the evolutionary mechanisms of the clustered regularly interspaced short palindromic repeat (CRISPR) locus in MTC.

1. Introduction

Despite the availability of effective antituberculosis chemotherapy for over 50 years (Styblo and Bumgarner, 1991), TB remains a major global health problem since the World Health Organization declared tuberculosis (TB) a global emergency in 1994 (Nakajima, 1993). The spread of multi-drug resistant tuberculosis (MDR-TB) and more recently of extremely drug resistant tuberculosis (XDR-TB) (Ralph et al., 2009), makes the implementation of public health measures, and molecular epidemiological investigations using rapid and high-throughput molecular methods an important point to follow TB transmission.

Current genotyping techniques used to study the epidemiology of *Mycobacterium tuberculosis* complex (MTC) clinical isolates are based on repetitive genetic elements: Clustered Repetitive Interspersed Short Palindromic Repeat (CRISPR) loci through the spoligotyping technique (Groenen et al., 1993; Kamerbeek et al., 1997; van Embden et al., 2000; Sorek et al., 2008) and Mycobacterial Interspersed Repetitive Units-Variable Number of Tandem repeats (MIRU-VNTR) (Frothingham and Meeker-O'Connell, 1998; Supply et al., 2000, 2001, 2006; Le Fleche et al., 2002; Skuce et al., 2002). Indeed, these markers have proven to be highly useful for epidemiological, population structural and evolutionary studies to distinguish between MTC clinical isolates (Abadia et al., 2009; Allix-Beguec et al., 2008a; Baranov et al., 2009; Brudey et al., 2006; Helal et al., 2009; Rohani et al., 2009; Stavrum et al., 2009) and have been used as the alternative to the classical IS*6110*-RFLP method (van Embden et al., 1993). These methods have in addition received recent technological improvements enabling fast and large-scale analyses to be performed (Cowan et al., 2004; Mazars et al., 2001). The classical spoligotyping procedure (Kamerbeek et al., 1997) relies on a reverse-line blot hybridization, a procedure that takes one full day of work to produce 43 profiles without interpretation; with the automatization now, 96 profiles can be

* Corresponding author at: Institute of Genetics and Microbiology, UMR8621, IGEPE Team, Universud, CNRS Université Paris-Sud 11, Campus d'Orsay, F-91405 Orsay Cedex, France; Tel.: +33 01 69 15 46 48; fax: +33 01 69 15 66 78.
E-mail address: christophe.sola@u-psud.fr (C. Sola).

1567-1348/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2010.07.006

obtained in half a day (Cowan et al., 2004; Zhang et al., 2010). This new technology relies on microbeads of different spectrum signatures to which capture probes are coupled depending on the targets (Cowan et al., 2004).

For phylogenetical evolutionary studies there is a concern using the fast evolving loci due to the presence of convergent evolutionary events within CRISPRs (Warren et al., 2002) as well as within MIRU-VNTR loci (Hanekom et al., 2008). To define phylogenetic associations unambiguously, genetic markers need to be unique and, ideally, irreversible (Comas et al., 2009).

In *M. tuberculosis*, these markers are large sequence polymorphisms (LSPs) (Hirsh et al., 2004; Mostowy et al., 2002) and single nucleotide polymorphisms (SNPs) (Alland et al., 2003; Dos Vultos et al., 2008; Filliol et al., 2006; Gutacker et al., 2006; Hershberg et al., 2008). LSPs are powerful markers in MTC because horizontal DNA transfer is extremely rare (Supply et al., 2003), but genetic distances based on genomic deletions are difficult to interpret in phylogenies (Gagneux et al., 2006). SNPs are less mutable than other forms of polymorphisms, making them unlikely to converge (Schork et al., 2000) so they are most appropriate markers for phylogenetic studies. In MTC strains structural genes exhibit rare polymorphism (Achtman, 2008; Kapur et al., 1994; Musser et al., 2000; Sreevatsan et al., 1997). However, recently Dos Vultos et al. (2008) have found higher polymorphisms in several genes involved in replication, recombination and repair functions (3R genes). For the first time, we provide a starting point of a new SNP typing of *M. tuberculosis* complex clinical isolates based mainly on 3R genes. We implemented this schedule on a high-throughput platform using a direct hybridization assay (Dunbar, 2006), that was used before for spoligotyping analyses (Cowan et al., 2004; Zhang et al., 2010).

We think that this could be the first step to provide a reliable high-throughput 3R SNP-based method for population structural studies and for further phylogenetical studies on *M. tuberculosis* complex clinical isolates.

2. Materials and methods

2.1. Chemicals, buffers and microbeads

All main materials and reagents required for microbead-based flow cytometry techniques were the same as described before (Zhang et al., 2010).

2.2. Oligonucleotides

To design capture probes and primers we used a demo of PrimerPlex (http://www.premierbiosoft.com/primerplex/index.html) and Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). All capture oligonucleotides (Eurogentec, Liège, Belgium) were manufactured and coupled as described before (Zhang et al., 2010). In this study we are targeting 7 genes that are polymorphic within *M. tuberculosis* sublineages, 6 of them from the 3R system described in Dos Vultos et al. (2008) and one (*mgtC*) reported by Alix et al. (2006). The primers and capture probes used here are listed in Tables 1 and 2, respectively.

2.3. Sequences analysis

Gene sequences were downloaded from http://genolist.pasteur.fr/TubercuList/. The DNAsp package (Rozas et al., 2003), Multalin (Corpet, 1988) and BioEdit sequence alignment editor (Hall, 1999) were used to locate SNPs on gene sequences and to align the gene and probe sequences in the design and/or verification of pre-designed capture probes.

Table 1
Genes, lineage or sublineage association, genomic position of targeted SNPs (up) and primers designed for the multiplex-PCR (down).

Gene[a]	Lineage or sublineage association	SNP targeted[b]
alkA	Bovine	807
recO	EAI	606
ligB	LAM	1212
recR	T2-related	94
mgtC	Haarlem	545
recC	X	1491
ligC	TUR-T3-Osaka	809

Primers

Gene	Sequence (5′–3′)	bp
alkA-F	CACGCTACGGCTCCCATG	18
alkA-R[c]	CCTTCGTCGATACCTGTTGGG	21
recO-F	TGTTGGACGCCTATCTGCTG	20
recO-R[c]	CCGTTCCAGATGCCATTGC	19
ligB-F	GGCTGGCTGAAGGTCAAG	18
ligB-R[c]	CATGGCGTCGGTCATTCC	18
recR-F	GGACCTGATTGACGAACTCG	20
recR-R[c]	GCCTGGATGTCTTTGGGTTC	20
mgtC-F	TCGTCGCTGTCCATCTCC	18
mgtC-R[c]	CACCAACCGCTCTAGCTTG	19
recC-F	CGCCGAACTGCTCACCATC	19
recC-R[c]	GCCACGCTTGGGAATCCTC	19
ligC-F	CGCGTCGGTCGGCGTGAT	18
ligC-R[c]	CGGGTCGCAGCGGCACGA	18

[a] Genes from the 3R system (Dos Vultos et al., 2008), except *mgtC* (Alix et al., 2006).
[b] nt position related to the gene.
[c] Reverse primers are biotin labeled.

2.4. SNPs typing PCR protocol

For direct hybridization of multiplexed-PCR assays, product length is recommended to be between 150 and 300 bp. We amplified segments around 200 bp of *alkA* (291 bp), *recO* (298 bp), *ligB* (174 bp), *recR* (255 bp), *mgtC* (272 bp) and *recC* (272 bp) to analyse the correlation between SNPs and major MTC lineages. To increase signals/cut-off ratio, PCR-multiplex was firstly divided into 3 sets, set1: *ligB*, *recR*, *mgtC* and *recC*; set2: *alkA* and *recO* and set3: *ligB* and *ligC*; however, running a single 7-Plex PCR protocol now provides similar results (see Supplemental Table 1). PCR assays were performed in 25 μL volumes of the following mixture: PCR buffer (100 mM Tris–HCl, pH 8.3, 500 mM KCl), 0.2 mM each dNTP, a final quantity of primers in a solution of 25 pmol of each one, 1.0 U Taq and 2 μL of DNA. The following PCR program was run: 5 min at 95 °C, followed by a touch-down PCR of 30 s at 95 °C,

Table 2
Sequence of the capture probes used to target the SNPs in this study.

Probes			
Probe name	Sequence 5′–3′	bp	Targeted site (nt)
alkA-wt	CGCGACCTGATGACGGC	17	807 ctg (leu) to
alkA-mut	CGCGACCTAATGACGGC	17	807 cta (leu)
recO-wt	GTACGACGGCGATTGGGA	18	606 ggc (gly) to
recO-mut	GTACGACGGTGATTGGGA	18	606 ggt (gly)
ligB-wt	GGCAAGCTCTCCAATATTCACC	22	1212 tcc (ser) to
ligB-mut	GGCAAGCTCTCGAATATTCACC	22	1212 tcg (ser)
RecR-wt	CTTCCACCTGTTGTCGGTAGA	21	94 ttg (leu) to
RecR-mut	CTTCCACCTGCTGTCGGTAGA	21	94 ctg (leu)
mgtC-wt	GGGGTATACGCACGGGGC	18	545 cgc (arg) to
mgtC-mut	GGGGTATACACACGGGGC	18	545 cac (his)
RecC-wt	GTGGCGGTTCGGACTCGA	18	1491 ttc (phe) to
RecC-mut	GTGGCGGTTGGGACTCGA	18	1491 ttg (leu)
ligC-wt	GACCACCCATGGAACTGGGCC	21	809 tgg (trp) to
ligC-mut	GACCACCCATTGAACTGGGCC	21	809 ttg (leu)

All capture probes have at 5′ a C-12 terminal linker. We are targeting both alleles, the wild type and the mutant for each site.

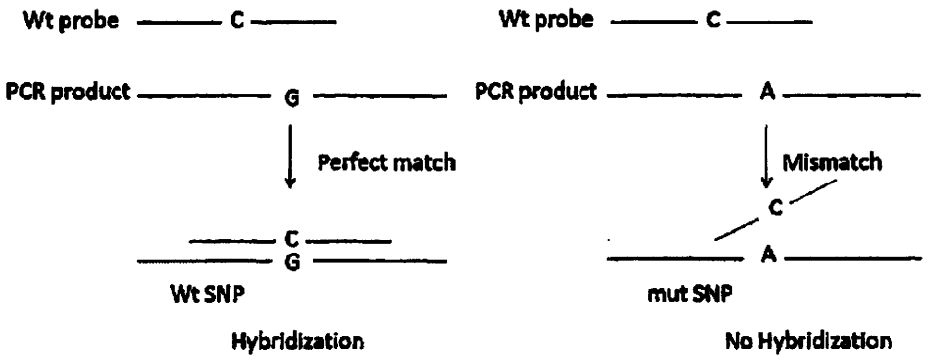

Fig. 1. Allele-specific hybridization. image modified from (Kwok, 2001).

30 s at 65 °C (−1 °C per step), 1 min at 72 °C 10 times and 30 s at 95 °C, 30 s at 56 °C, 1 min at 72 °C for 29 more cycles with a final extension step of 10 min at 72 °C. The PCR amplifications were checked during first tests by agarose gel electrophoresis.

2.5. Hybridization

The procedure detecting the two possible alleles relies on an allele-specific hybridization (Kwok, 2001). Fig. 1 shows how the SNPs were discriminated based on this approach. If the PCR product harbors the wild type (wt) SNP it will match to the wt probe that is complementary to the sequence. This hybridized product is very stable because there is a 100% complementarity. If the PCR product has the mutated (mut) SNP allele, the hybridization is not thermodynamically stable due to the nucleotide mismatch and it will be lost. For this kind of procedure the mismatch base should be placed in the middle of the probe. The hybridization procedure on microbeads was done as described previously (Zhang et al., 2010) except for the second part of the hybridization that was performed at 52 °C for 20 min.

2.6. Mycobacterium isolates, origin, DNA extraction

327 DNA samples (a summary of the lineages being targeted can be found in Supplemental Table 2) were genotyped by our novel SNP typing scheme using a high-throughput microbead-based method, as reported previously (Cowan et al., 2004; Zhang et al., 2010). The clinical isolates for this study were chosen to cover the main lineages of *M. tuberculosis* complex clinical isolates described before by spoligotyping (Filliol et al., 2002; Brudey et al., 2006). They were selected from an international quality control study on membrane-based spoligotyping vs. high-throughput-based spoligotyping (Abadia et al., unpublished results) as well as from the 2004–2008 collection of the TB National Reference Laboratory in the Netherlands, from Zonguldak hospital, Turkey and from a Medical Center in Japan. More specifically, 59 samples were from Buenos Aires – Argentina (Servicio de Micobacterias, Instituto Nacional de Enfermedades Infecciosas, ANLIS "Carlos G. Malbran", Buenos Aires, Argentina), 109 samples came from Bilthoven – The Netherlands (National Institute for Public Health and the Environment – RIVM), 120 DNA samples from the Faculty of Medicine, Department of Microbiology and Clinical Microbiology, Zonguldak Karaelmas University (Zonguldak – Turkey), 40 DNA samples from Japan (Osaka Prefectural Medical Center for Respiratory and Allergic Diseases). The DNA from these samples was extracted either by a simple thermolyzate or by the classical cetyltrimethylammonium bromide (CTAB) procedure.

2.7. Data analysis

Standard MTC lineage assignation (*i.e.* according to spoligotype pattern) was reported either according to spolDB4 database (Brudey et al., 2006) or using the SPOTCLUST algorithm (Vitol et al., 2006) online access http://cgi2.cs.rpi.edu/~bennek/Run.html).

To infer lineages according to SNPs, the presence of allele putative-lineages associated (Dos Vultos et al., 2008) was reported. Hybridization signals for the mutated or wild type SNP alleles, collected as RFI (Relative Fluorescence Intensity) were transformed in a binary code (presence/absence of each allele) using a signal/noise cut-off value of 2.

3. Results

The aim of this study was: (1) to test some associations between spoligotype-defined lineages and a set of potentially lineage-specific SNPs mainly located on 3R genes (Alix et al., 2006; Dos Vultos et al., 2008) on a representative set of samples; (2) to develop a new multiplexed high-throughput assay for this purpose; (3) to assign a lineage to those clinical isolates for which spoligotype signatures were uninformative, the so-called "U" clinical isolates in SpolDB4 (Brudey et al., 2006); and finally (4) to confirm previous lineage associations or correct possible mis-assignations. We thus aimed to further contribute to the study of the MTC molecular evolution.

We searched on the 3R gene database (Dos Vultos et al., 2008) for SNPs that were polymorphic between main MTC lineages and we found that SNPs on: *alkA* (807 ctg [leu] to cta [leu]), *recO* (606 ggc [gly] to ggt [gly]), *ligB* (1212 tcc [ser] to tcg [ser]), *recC* (1491 ttc [phe] to ttg [leu]), *recR* (94 ttg [leu] ctg [leu]), *ligC* (809 tgg [trp] to ttg [leu]) apparently were associated with *M. bovis*, East African (EAI), the Latin-American Mediterranean (LAM), the X, the T2 and the LAM7_TUR/T3_Osaka *M. tuberculosis* lineages, respectively. We added to this set of genes the SNP on *mgtC* 545 cgc [arg]) to cac [his] because it was described previously as being associated with the Haarlem lineage (Alix et al., 2006) and recently had been tested by Chuang et al. (2008). The mgtC protein is a common virulence factor to several intracellular pathogens (Buchmeier et al., 2000).

Our scheme thus targets 7 lineages: 5 large lineages covering a total of 31 sublineages, and two specific sublineages. Other SNPs belonging to other panel of genes could also be added in a near future to our assay. We have developed a single multiplexing reaction including this set of 7 genes to test the previous finding of 3R SNP MTC lineage association in a representative set of samples. Central-Asian (CAS) specific and East-Asian (Beijing) specific SNPs are not reported at this stage of work but are also in progress.

To validate our new SNP typing method, we used 327 DNA samples of *M. tuberculosis* clinical isolates, all having been previously spoligotyped on microbeads using the high-throughput technique (Zhang et al., 2010). These samples were representative of a highly diverse MTC population within the total worldwide genetic diversity of MTC: they covered unclassified strains, as well as 32 sublineages over the 62 that have been named in SpolDB4, representing all main lineages infecting humans including some that were not targeted by our SNP scheme (Beijing, CAS, EAI, LAM, bovis, Haarlem, X, S and T). Only *Mycobacterium africanum* and *Mycobacterium canettii* and Manu lineages were not represented. Out of the 32 sublineages that were represented in our data set, 21 were targeted by our scheme.

Among the seven mutations tested within our scheme we never detected two mutations simultaneously in any of the 327 typed strains, leading either to SNP-specific assignment or to SNP-nonspecific assignment. In addition, none of the isolates assigned to lineages Beijing, CAS, S and T4 did show any of the mutation present in our SNP scheme as expected. This highlights the specificity of the SNP we chose, and confirms the absence of convergence events in this collection. For sixteen sublineages (76%, covering 73 different isolates) targeted by our scheme and present in our data set, complete congruence between spoligotyping-based classification and SNP-based assignation was observed, indicating a high global fit between the two classifications: 6/6 *M. bovis*, 55 EAI (EAI1_SOM, EAI2_MANILLA, EAI5, EAI3), 28/38 H and 14/14 from sublineages H1 and H2, 23/37 LAM and 23/23 from sublineages LAM2, LAM3, LAM5, LAM9, 2/2 X (X1, X3), 23/23 T3-Osaka. For five sublineages that were targeted by the scheme, part or all of the assignations were different between the two classifications: 4/18 clinical isolates from H3, 6/6 from H4, 28/34 from T2, 7/7 from LAM3-S convergent, 40/40 LAM7_TUR (Table 3 for an overview; spoligotype patterns and SNP genotypes in Supplemental Table 3). One T5 sublineage (T5_RUS1) and 6 T1-Tuscany, and one T1 however carried the LAM-specific SNP, 2 other T1 isolates carried the H-specific mutation, and 2 other T1 isolates the T3-Osaka specific SNP, confirming that T-clade is poorly defined.

More specifically, not all of the clinical isolates labeled as Haarlem lineage (H1 to H4) according to their spoligotype signatures in SpolDB4 ($n = 38$) harbored the SNP Haarlem associated on *mgtC*$^{545\ (cgc\text{-}cac)}$ (Alix et al., 2006). All H1 spoligo-signatures ($n = 11$) and all H2 spoligo-signatures ($n = 3$) had this specific mutation whereas 14/18 of H3 spoligo-signature and none of the H4 ($n = 6$) had. Among the four H3 clinical isolates that did not harbor the *mgtC* mutation, two belonged to SIT316, one to

Table 3
SNP lineage assignation on a subset of analyzed strains and comparison with SpolDB4 linage classification to show in some cases the lineage confirmation and in others a lineage re-assignation.

		Lineage[a]		
SIT	Spoligotype	SpolDB4	SNP	No. of strains[e]
35		H4	-	2
262		H4	-	3
777		H4	-	1
760		H3	-	1
134		H3	X	1
316		H3	T2	2
47[b]		H1	H	1
50[b]		H3	H	9
41		LAM7_TUR	TUR-T3-Osa	6
367		LAM7_TUR	TUR-T3-Osa	1
4		LAM3-S/conv	-	7
42[b]		LAM9	LAM	10
33[b]		LAM3	LAM	4
93[b]		LAM5	LAM	1
17[b]		LAM2	LAM	2
52		T2	T2[c]	31[c]
254		T5_RUS1	LAM	1
736		T2	-	1
627		T3-OSA	TUR-T3-Osa	40
78		T1-T2	TUR-T3-Osa	2
53**		T1	TUR-T3-Osa[d]	25[d]
159		T1 (Tuscany v.)	LAM	7
370		T1	LAM	1
53		T1	H	1
120		T1	H	1

[a]Family assignation according to spoligo-signatures and according to SNP identification.
[b]Reference strains for each lineage.
[c]Just 6/31 had the mutation on *recR*$^{94\ (ag\text{-}cg)}$ T2-related.
[d]2/25 strains SIT 53 had the *ligC*^{809} SNP.
[e]Number of strains related to the frequency of the spoligotype pattern.
To read and interpret the spoligotype, spacers are placed in order from 1 to 43 (left to right).
"-": Means no detectable SNP.

SIT760 and the other to SIT134. The SIT134 "H3" harbored the *recC*$^{1491\ (ttc\text{-}ttg)}$ X lineage associated SNP and the two SIT316 "H3" clinical isolates harbored the mutation on *recR*$^{94\ (ttg\text{-}ctg)}$ hypothesized to be linked to the T2 lineage isolates from Central African Republic clinical isolates (Dos Vultos et al., 2008) (see also Table 3 and Supplemental Table 3).

Within LAM-labeled clinical isolates ($n = 72$), we found that 47 did not harbor the expected mutation on *ligB*$^{1212\ (tcc\text{-}tcg)}$, forty of these were LAM7_TUR and the seven others were LAM3-S (Table 3). All others LAM sublineages tested so far (LAM9, LAM5, LAM3, and LAM2) did harbor the expected *ligB* mutation.

Within the T lineage, there are a lot of sublineages defined by spoligotype signatures (Brudey et al., 2006), among which one is T2. *recR*$^{94\ (ttg\text{-}ctg)}$ SNP was found associated with Central African Republic isolates, all of them carrying a T2 signature (absence of spacer 40) (Dos Vultos et al., 2008). Here, 83 DNA samples belonged to the T lineage (sublineages: either T1, T1-Tuscany variant, T1–T2 undefined, T2, T4_CEU1, T5 and T5_RUS1), out of 32 T2 isolates, six only harbored the mutation on *recR*$^{94\ (ttg\text{-}ctg)}$ (Table 3). These results show that the *recR*^{94} mutation cannot be considered as being linked to the absence of spacer 40, which defines the T2 sublineage.

Table 4
Lineage identification of spoligotype that lack of it, the so-called "U" spoligotypes in SpolDB4 (Brudey et al., 2006), either because they were not reported in the database, or because this identification was not clear. Lineage assignment by SNP and by SPOTCLUST (Vitol et al., 2006) is shown.

			Family			
Pos[a]	SIT	Spoligotype	SpolDB4	SNP	No. of strains[c]	SPOTCLUST[b]
1	NR		U	EAI	2	EAI5 (0.95)
2	NR		U	EAI	1	EAI5 (1.00)
3	NR		U	H	1	Family 34 (1.00)
4	NR		U	H	3	Family 35 (1.00)
5	NR		U	H	1	H1 (1.00)
6	NR		U	H	2	H3 (0.77)
7	NR		U	H	1	H3 (0.96)
8	1274		U	H	1	H1 (1.00)
9	NR		U	H	1	H3 (0.77)
10	NR		U	H	1	H1 (1.00)
11	NR		U	H	1	H1 (1.00)
12	NR		U	H	2	H3 (0.77)
13	NR		U	H	1	H1 (1.00)
14	NR		U	LAM	2	T1 (0.80)
15	NR		U	LAM	1	LAM3 (0.93)
16	NR		U	LAM	4	LAM9 (1.00)
17	NR		U	LAM	1	LAM8 (0.94)
18	NR		U	LAM	1	LAM3 (1.00)
19	NR		U	LAM	1	LAM8 (0.99)
20	NR		U	LAM	1	LAM3 (1.00)
21	NR		U	LAM	1	LAM9 (0.96)
22	NR		U	LAM	1	LAM9 (1.00)
23	NR		U	LAM	1	T4 (1.00)
24	NR		U	LAM	2	LAM (1.00)
25	105		U (LAM3?)	LAM	2	LAM7 (1.00)
26	NR		U	LAM	1	LAM8 (1.00)
27	NR		U	LAM	1	LAM9 (0.97)
28	NR		U	LAM	1	T4 (1.00)
29	1531		U	X	1	X1 (0.65)
30	NR		U	X	2	X3 (1.00)
31	NR		U	X	1	H3 (0.77)
32	2125[d]		U	TUR-T3-Osa	1	LAM9 (1.00)
33	NR		U	TUR-T3-Osa	1	LAM9 (0.99)
34	NR		U	TUR-T3-Osa	1	T1 (1.00)
35	NR		U	TUR-T3-Osa	1	T1 (1.00)

[a]This number is provided just to locate the pattern in this list.
[b]SPOTCLUST Identification according the spoligotype pattern and the algorithm probability of the result of being true in parenthesis.
[c]Number of strains related to the frequency of the spoligotype pattern.
[d]SIT number from Aktas et al. (2008).
To read and interpret the spoligotype, spacers are placed in order from 1 to 43 (left to right).

An interesting secondary result found in this study is the confirmation of the phylogenetic link between the clinical isolates from the T3_Osaka sublineage and the ones from the LAM7_TUR. These two groups of strains share the same MIRU-VNTR[1] 12 loci signature (MIRU-VNTR International Type, MIT310) and harbor divergent spoligotypes (SIT41 and SIT627). Here, using a large set of DNA samples from Zonguldak – Turkey (LAM7_TUR) and from Japan (T3_Osaka and other variants), our results show that they shared the same SNP on *ligC*$^{809\ (tgg\ [trp]\ to\ ttg\ [leu])}$. Moreover, we observed that 2/2 SIT78 (T1–T2) and 2/25 SIT53 (T1) also had this SNP.

Our study also included 84 spoligotypes for which the lineage information was unknown, the so-called "U" isolates (for Unknown), *i.e.* isolates for which the spoligotype signature either did not allow assignation to a lineage ($n = 7$) (Brudey et al., 2006) or had not been reported in the international spoligotype database, SpolDB4 ($n = 77$, referred here as "NR"). With our SNP-based scheme, 47 clinical isolates (56%, with same frequency for strict "U" and "NR") could now be assigned at a lineage level (Table 4). The identification of these clinical isolates is distributed as follows: 21 isolates were assigned to the LAM-lineage, 15 to Haarlem, 4 to X and 3 to the EAI lineage. We also tested the congruence between our SNP classification and SPOTCLUST, a Probabilistic Multivariate Bernoulli Mixture Model that uses Naïve Bayes assumptions, as reported previously (Vitol et al., 2006). This algorithm automatically provides a probability for a spoligotype pattern to be part of a lineage using previous assignations reported in SpolDB4 (Brudey et al., 2006). Interestingly, we found a good correlation (72%) between SPOTCLUST assignation and our SNP identification (Table 4). However, we also observed discrepancies (positions 3, 4, 14, 23, 28 and 31 in Table 4). Indeed, SPOTCLUST classified these clinical isolates as: Lineage 34, Lineage 35, T1, T4, T4 and H3, respectively with probabilities of higher than 0.77 whereas our results classify these samples as H, H, LAM, LAM, LAM and X, respectively. Note that TUR_T3_Osaka strains were identified by SPOTCLUST as LAM9 or T1 with probabilities close to 1.00 (Table 4).

4. Discussion

We aimed to confirm the spoligotype-based classification of MTC using a restricted SNP-based scheme designed to be performed in a high-throughput way. Most commonly used classification of MTC relies on spoligotype patterns. Lineages and sublineages have been defined based on recurrent so-called spoligo-signatures (absence and/or presence of specific spacers) that have proven to be phylogeographically meaningful (Brudey et al., 2006). The existence of these lineages was confirmed using independent techniques such as LSP (Gagneux et al., 2006), MIRU-VNTR typing (Allix-Beguec et al., 2008b; Wirth et al., 2008)), or via sequencing projects (Comas et al., 2009).

To provide phylogenetically reliable information, markers should not be prone to converge. SNPs most of the time are unlikely to convergence and also are less prone to distortion by selective pressures (Schork et al., 2000). Notwithstanding drug resistance genes, there is just one evidence of convergence evolution due to a nonsense variation in ada/alkA gene observed in *M. tuberculosis* and in *M. bovis* that could confer a selective advantage (Nouvel et al., 2007), a SNP that is not included in our data set.

In *M. tuberculosis*, and maybe other clonal organisms, the genes involved in replication, repair and recombination (3R) seemed to play a key role in adaptation (Dos Vultos et al., 2008) so they might be prone to experience selection. However, in our study, four of the targeted SNPs are synonymous (sSNPs) so they are unlikely to be selected for. In addition, the three other non-synonymous mutations chosen are conservative so the replaced amino acid is biochemically similar to the pre-existing one. Altogether, the SNPs we provide are likely to be neutral. Regarding classification, we confirmed that EAI, BOVIS and X lineages are monophyletic, *i.e.* that the signature that defines them is phylogenetically relevant: EAI lineage is defined by the absence of spacers 29–32, absence of 34 and the *recO*$^{606\ (ggc\text{-}ggt)}$ mutation; the BOVIS lineage is both defined by the absence of spacers 39–43 and the *alkA*$^{807\ (ctg\text{-}cta)}$ mutation; the X lineage is defined by the absence of spacer 18, 33–36 group of spacers and by the *recC*$^{1491\ (tcc\text{-}ttg)}$ mutation. The results we provide come in addition to previous results obtained (Dos Vultos et al., 2008) so that we can say that there is strong support for these correlations.

This study in contrast provides some evidence of mis-labeling in LAM, T and Haarlem lineages due to the used marker itself. Regarding Haarlem lineage, H4 sublineage did not carry the expected *mgtC*$^{545\ (cgc\text{-}cac)}$ mutation. Kovalev et al. (2005) showed that according to their MIRU-VNTR signatures (especially MIRU10 with 7–10 copies) this lineage was unlikely to be phylogenetically related to Haarlem so it was renamed as URAL. Here, we confirm that the loss of spacer 29–31 in URAL is a genetic event that is independent from the loss of spacer 26–31 in H1 and H2, or of spacer 31 in H3. For the time being we have not identified a SNP that correlates with this lineage. Interestingly, all the strains that carried this signature in our scheme came from Turkey that is geographically close to Russia.

Two SITs (134 and 316) that belong to the H3 lineage (according to their spoligotype signatures) lacked the *mgtC*^{545} SNP Haarlem-specific mutation but instead harbored the *recC*^{1491}-X and *recR*^{94}-T2 lineage-specific mutations, respectively that may indicate that the deletion of spacer 31 could have occurred several times independently so that it could be used cautiously as a phylogenetic marker. We suggest that a "Haarlem" lineage assignment for strains classified as H3 could therefore be confirmed by the SNP method described here.

The X lineage was first defined by the absence of spacer 18 (Sebban et al., 2002; Filliol et al., 2002) in addition to the absence of spacers 33–36. Here, we found a strain that was not labeled as X according to its spoligotype, but in fact it was according to its SNP pattern. This suggests that the deletion of spacer 18 occurred after the *recC*^{1491} SNP event. Hence, we suggest to broaden X lineage to all strains carrying this SNP.

Within the T2 sublineage (defined by the absence of spacer 40, in addition to the absence of 33–36 group of spacers) many clinical isolates sharing SIT52 or SIT736 spoligotype did not harbor the SNP on *recR*^{94}. In addition a relatively low number of *recR*^{94} mutations were found in the whole set of strains, suggesting that this mutation occurred too recently. For the time being, we are not able to well describe the T lineage and resolve all the relationships between modern strains carrying the 33–36 spacer deletion, because there is no single SNP available to do that by now and due to the big genetic diversity inside this main lineage, it remains "ill-defined" or poorly defined (Brudey et al., 2006). Also, Vitol et al. (2006) were unable to create models that discriminate well among the members of the T lineage based on their spoligotype signatures.

Another contribution to the T sublineage markers is the confirmation (with strong support, $n = 62$) of the existence of a TUR_T3_Osaka lineage based on the *ligC*^{809} SNP. Spoligotypes that harbors this SNP are very different between them (Table 3 and Supplementary Table 1). Millet et al. (2007) reported before a shared MIRU12-VNTR pattern called MIT310 (215125113322) between strains from T3_Osaka lineage (Takashima and Iwamoto, 2006) and the previously reported LAM7_TUR lineage (Zozio et al., 2005; Millet et al., 2007). All the LAM7_TUR strains tested so far lacked the mutation on the LAM-associated *ligB*^{1212} SNP, consequently they are

[1] Mycobacterial Interspersed Repetitive Unit-Variable Number of Tandem Repeat.

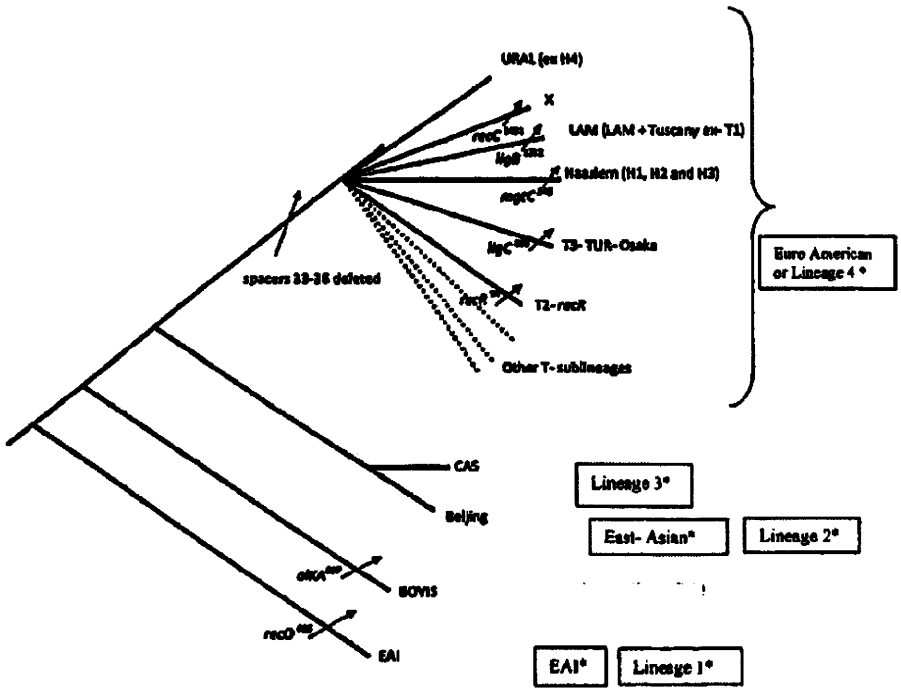

Fig. 2. Evolutionary proposal to identify *M. tuberculosis* families based on the SNP used in this study. Small arrows show the genes in which the SNP is present to allow the identification. The big arrow shows the deletion of spacers 33–36 that is common to all the modern strains. It is shown the correspondence between major lineages defined by Spoligotyping, our SNP proposal and inside boxes the one related to LSP plus a recently work done in 89 sequenced genes. *The LSP classification is related to the work of (Gagneux et al., 2006) and recently to an SNP work of (Comas et al., 2009).

not a sublineage of LAM so we give further support to renaming them TUR (Table 3). We also observed that strains carrying SIT78 T1–T2, SIT53 T1 (2/25), or SIT553 T2–T3 and 4 NR strains also harbored the TUR-T3-Osaka-*ligC*^{809} mutation. This suggests that this mutation could be useful digging inside a scheme targeting T_lineage genetic diversity.

Another result is that one SIT53 T1 strain, which is both wt for *recR*^{94} and for *ligC*^{809}, harbors the Haarlem-specific mutation for *mgtC*^{545} (Table 3). This finding could corroborate the hypothesis made before by others (Duchene et al., 2004) about the relation between T and Haarlem strains. In this case, one possibility is that strains harboring the SIT53 T1 spoligo-signature may have evolved at a genome level to Haarlem in spite of retaining the CRISPR region of the progenitor. This could be also the case for SIT120 T1 for which we found that it also harbors the mutation for Haarlem on *mgtC*^{545}, the only difference at a spoligo level between SIT53 and 120 is one spacer missing, spacer 20 (Table 3).

With our SNP-signature panel we also provide evidence favoring the renaming of several spoligotypes previously assigned to T-clade based on spoligo rules, these are: SIT159 T1 (Tuscany variant), SIT254 (T5_RUS1) and SIT370 (T1) that harbored the *ligB*^{1212} SNP-LAM-associated mutation so we assigned the right lineage to these strains. The LAM assignation rule is based on the simultaneous absence of spacers 21–24 and 33–36 that is not accomplished by SIT159, SIT 254 T5_RUS1or SIT370. Taken together, this finding could be an indication that spacers 21 and 22 are not so informative to define the LAM lineage as is the absence of spacer 23.

One of the most important features of our new SNP proposal relies on the identification of those spoligotypes that lack the lineage identification. Emerging clones continuously will appear as a result of genetic variation of pre-existing ones (Brudey et al., 2006). The SNP assignment is unambiguous, providing a precise tool that targets several sites (7 in our actual proposal) with a high precision performing the lineage assignation. We could successfully assign a lineage to 47 out of 84 U strains. One interesting lineage assignment in the group of "U" strains was the SIT 105 that had been responsible for an MDR-TB outbreak in Spain. This SIT was thought to be "LAM3?" in SpolDB4 with doubts and now we could confirm it because it was found to harbor the *ligB*^{1212}-LAM-specific mutation.

Even though Spoligotyping remains a first line tool to delineate the molecular ecology of the circulating strains, we have shown that sometimes the spoligotype signature will not always reflect the real lineage of the MTC strains. With this work, we have solved some intraspecies taxonomic issues and we provide an accurate lineage assignation using a minimal set of SNPs almost as informative as large sequencing projects similar to those of Comas et al. (2009). However, we still need to include other SNPs for other lineages, because there are still spoligotype patterns that we were not able to identify with this proposal. This is also true to reach the level of spoligotyping sublineage discrimination. Nevertheless we could correctly identify strains belonging to the major MTC lineages or families such as EAI, BOVIS, X, LAM, Haarlem as well as some T sublineages, such as T2-*recR*^{94} and TUR_T3_OSAKA. We do not have one specific SNP for the T or S strain lineage for the time being. In Fig. 2 we show an evolutionary scenario built with the set of SNPs used in this study and a hypothesis of when the deletion of spacers 33–36 may have occurred.

Some SNP-based studies have already been performed in *M. tuberculosis* to date (Alland et al., 2003; Baker et al., 2004; Filliol et al., 2006; Gutacker et al., 2006; Gutacker et al., 2002; Sreevatsan et al., 1997). However, the composition and number of SNP cluster groups (SCGs) within *M. tuberculosis* have remained unclear since it also used some inadequate genes and SNPs (i.e. involved in drug resistance) (Baker et al., 2004) whereas others were selected from a non-representative set of available genomes (Filliol et al., 2006; Gutacker et al., 2002, 2006). These previous studies had not taken full advantage of the power of SNP-based methods. Here we show that SNP typing could serve as a "gold standard" for DNA typing. We think that in the near future the inclusion of more SNP information will become an important parameter to efficiently

classify a given MTC clone, either for molecular epidemiological or for evolutionary purposes. High-throughput multiplexing among various techniques and platforms will be one of the best ways to achieve this with reasonable economic costs (Dunbar, 2006; Bergval et al., 2008).

Acknowledgments

This study was supported by the "Excellency Chair in microbiology of the Université Paris-Sud 11" granted to CS in September 2007, through a 12-month post-doctoral fellowship grant given to EA. EA is a senior mycobacteriology scientist of the IVIC (Instituto Venezolano de Investigaciones Científicas, Caracas, Venezuela). JZ is a PhD fellow of the Université Paris-Sud 11 through the "Ecole Doctorale Gènes Génomes Cellules". We are grateful to François Topin, Luminex BV, The Netherlands, and to Luminex Corporation, Austin, Texas for their support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.meegid.2010.07.006.

References

Abadia, E., Sequera, M., Ortega, D., Mendez, M.V., Escalona, A., Da Mata, O., Izarra, E., Rojas, Y., Jaspe, R., Motiwala, A.S., Alland, D., de Waard, J., Takiff, H.E., 2009. *Mycobacterium tuberculosis* ecology in Venezuela: epidemiologic correlates of common spoligotypes and a large clonal cluster defined by MIRU-VNTR-24. BMC Infect. Dis. 9, 122.

Achtman, M., 2008. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu. Rev. Microbiol. 62, 53–70.

Aktas, E., Zozio, T., Comert, F.B., Kulah, C., Aydin, O., Rastogi, N., Sola, C., 2008. A first insight into the genetic diversity and population structure of *Mycobacterium tuberculosis* in Zonguldak, Turkey. Clin. Microbiol. Infect. 55–59.

Alix, E., Godreuil, S., Blanc-Potard, A.B., 2006. Identification of a Haarlem genotype-specific single nucleotide polymorphism in the mgtC virulence gene of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 44, 2093–2098.

Alland, D., Whittam, T.S., Murray, M.B., Cave, M.D., Hazbon, M.H., Dix, K., Kokoris, M., Duesterhoeft, A., Eisen, J.A., Fraser, C.M., Fleischmann, R.D., 2003. Modeling bacterial evolution with comparative-genome-based marker systems: application to *Mycobacterium tuberculosis* evolution and pathogenesis. J. Bacteriol. 185, 3392–3399.

Allix-Beguec, C., Fauville-Dufaux, M., Supply, P., 2008a. Three-year population-based evaluation of standardized mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 46, 1398–1406.

Allix-Beguec, C., Harmsen, D., Weniger, T., Supply, P., Niemann, S., 2008b. Evaluation and strategy for use of MIRU-VNTRplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of *Mycobacterium tuberculosis* complex isolates. J. Clin. Microbiol. 46, 2692–2699.

Baker, L., Brown, T., Maiden, M.C., Drobniewski, F., 2004. Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*. Emerg. Infect. Dis 10, 1568–1577.

Baranov, A.A., Mariandyshev, A.O., Mannsaker, T., Dahle, U.R., Bjune, G.A., 2009. Molecular epidemiology and drug resistance of widespread genotypes of *Mycobacterium tuberculosis* in northwestern Russia. Int. J. Tuberc. Lung Dis. 13, 1288–1293.

Bergval, I.L., Vijzelaar, R.N., Dalla Costa, E.R., Schultema, A.R., Oskam, L., Kritski, A.L., Klatser, P.R., Anthony, R.M., 2008. Development of multiplex assay for rapid characterization of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 46, 689–699.

Brudey, K., Driscoll, J., Rigouts, L., Prodinger, W.M., Gori, A., Al-Hajoj, S.A.M., Allix, C., Aristimuno, L., Arora, J., Baumanis, V., Binder, L., Cafrune, P., Cataldi, A., Cheong, S., Diel, R., Ellermeier, C., Evans, J.T., Fauville-Dufaux, M., Ferdinand, S., Garcia de Viedma, D., Garzelli, C., Gazzola, L., Gomes, H.M., Gutierrez, M.C., Hawkey, P.M., van Helden, P.D., Kadival, G.V., Kreiswirth, B.N., Kremer, K., Kubin, M., Kulkarni, S.P., Liens, B., Lillebaek, T., Ly, H.M., Martin, C., Martin, C., Mokrousov, I., Narvskaia, O., Ngeow, Y.F., Naumann, L., Niemann, S., Parwati, I., Rahim, M.Z., Rasolofo-Razanamparany, V., Rasolonavalona, T., Rossetti, M.L., Rüsch-Gerdes, S., Sajduda, A., Samper, S., Shemyakin, I.G., Singh, U.B., Somoskovi, A., Skuce, R., Van Soolingen, D., Streicher, E.M., Suffys, P.N., Tortoli, E., Tracevska, T., Vincent, V., Victor, T.C., Warren, R., Yap, S.F., Zaman, K., Portaels, F., Rastogi, N., Sola, C., 2006. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics, and epidemiology. BMC Microbiol. 6, 23.

Buchmeier, N., Blanc-Potard, A., Ehrt, S., Piddington, D., Riley, L., Groisman, E.A., 2000. A parallel intraphagosomal survival strategy shared by *Mycobacterium tuberculosis* and *Salmonella enterica*. Mol. Microbiol. 35, 1375–1382.

Chuang, P.C., Liu, H., Sola, C., Chen, Y.M., Jou, R., 2008. Spoligotypes of *Mycobacterium tuberculosis* isolates of a high tuberculosis burden aboriginal township in Taiwan. Infect. Genet. Evol. 8, 553–557.

Comas, I., Homolka, S., Niemann, S., Gagneux, S., 2009. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. PLoS ONE 4, e7815.

Corpet, F., 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. 16, 10881–10890.

Cowan, L.S., Diem, L., Brake, M.C., Crawford, J.T., 2004. Transfer of a *Mycobacterium tuberculosis* genotyping method, Spoligotyping, from a reverse line-blot hybridization, membrane-based assay to the Luminex multianalyte profiling system. J. Clin. Microbiol. 42, 474–477.

Dos Vultos, T., Mestre, O., Rauzier, J., Golec, M., Rastogi, N., Rasolofo, V., Tonjum, T., Sola, C., Matic, I., Gicquel, B., 2008. Evolution and diversity of clonal bacteria: the paradigm of *Mycobacterium tuberculosis*. PLoS ONE 3, e1538.

Duchene, V., Ferdinand, S., Filliol, I., Guégan, J.F., Rastogi, N., Sola, C., 2004. Phylogenetic reconstruction of the *Mycobacterium tuberculosis* complex within four settings of the Caribbean region: tree comparative analysis and first appraisal on their phylogeography. Infect. Gen. Evol. 4, 5–14.

Dunbar, S.A., 2006. Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection. Clin. Chim. Acta 363, 71–82.

Filliol, I., Driscoll, J.R., Van Soolingen, D., Kreiswirth, B.N., Kremer, K., Valetudie, G., Anh, D.D., Barlow, R., Banerjee, D., Bifani, P.J., Brudey, K., Cataldi, A., Cooksey, R.C., Cousins, D.V., Dale, J.W., Dellagostin, O.A., Drobniewski, F., Engelmann, G., Ferdinand, S., Gascoyne-Binzi, D., Gordon, M., Gutierrez, M.C., Haas, W.H., Heersma, H., Kallenius, G., Kassa-Kelembho, E., Koivula, T., Ly, H.M., Makristathis, A., Mammina, C., Martin, G., Mostrom, P., Mokrousov, I., Narbonne, V., Narvskaya, O., Nastasi, A., Niobe-Eyangoh, S.N., Pape, J.W., Rasolofo-Razanamparany, V., Ridell, M., Rossetti, M.L., Stauffer, F., Suffys, P.N., Takiff, H., Texier-Maugein, J., Vincent, V., De Waard, J.H., Sola, C., Rastogi, N., 2002. Global distribution of *Mycobacterium tuberculosis* spoligotypes. Emerg. Infect. Dis. 8, 1347–1349.

Filliol, I., Motiwala, A.S., Cavatore, M., Qi, W., Hernando Hazbon, M., Bobadilla Del Valle, M., Fyfe, J., Garcia-Garcia, L., Rastogi, N., Sola, C., Zozio, T., Guerrero, M.I., Leon, C.I., Crabtree, J., Angiuoli, S., Eisenach, K.D., Durmaz, R., Joloba, M.L., Rendon, A., Sifuentes-Osornio, J., Ponce de Leon, A., Cave, M.D., Fleischmann, R., Whittam, T.S., Alland, D., 2006. Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. J. Bacteriol. 188, 759–772.

Frothingham, R., Meeker-O'Connell, W.A., 1998. Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats. Microbiology 144 (Pt 5), 1189–1196.

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B.C., Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M.C., Hilty, M., Hopewell, P.C., Small, P.M., 2006. Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. U.S.A. 103, 2869–2873.

Groenen, P.M., Bunschoten, A.E., van Soolingen, D., van Embden, J.D., 1993. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. Mol. Microbiol. 10, 1057–1065.

Gutacker, M.M., Mathema, B., Soini, H., Shashkina, E., Kreiswirth, B.N., Graviss, E.A., Musser, J.M., 2006. Single-nucleotide polymorphism-based population genetic analysis of *Mycobacterium tuberculosis* strains from 4 geographic sites. J. Infect. Dis. 193, 121–128.

Gutacker, M.M., Smoot, J.C., Migliaccio, C.A., Ricklefs, S.M., Hua, S., Cousins, D.V., Graviss, E.A., Shashkina, E., Kreiswirth, B.N., Musser, J.M., 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.

Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for indows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Hanekom, M., van der Spuy, G.D., Gey van Pittius, N.C., McEvoy, C.R., Hoek, K.G., Ndabambi, S.L., Jordaan, A.M., Victor, T.C., van Helden, P.D., Warren, R.M., 2008. Discordance between mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing and IS6110 restriction fragment length polymorphism genotyping for analysis of *Mycobacterium tuberculosis* Beijing strains in a setting of high incidence of tuberculosis. J. Clin. Microbiol. 46, 3338–3345.

Helal, Z.H., Ashour, M.S., Eissa, S.A., Abd-Elatef, G., Zozio, T., Babapoor, S., Rastogi, N., Khan, M.I., 2009. Unexpectedly high proportion of ancestral Manu genotype *Mycobacterium tuberculosis* strains cultured from tuberculosis patients in Egypt. J. Clin. Microbiol. 47, 2794–2801.

Hershberg, R., Lipatov, M., Small, P.M., Sheffer, H., Niemann, S., Homolka, S., Roach, J.C., Kremer, K., Petrov, D.A., Feldman, M.W., Gagneux, S., 2008. High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. PLoS Biol. 6, e311.

Hirsh, A.E., Tsolaki, A.G., DeRiemer, K., Feldman, M.W., Small, P.M., 2004. Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. Proc. Natl. Acad. Sci. U.S.A. 101, 4871–4876.

Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., Bunschoten, A., Molhuizen, H., Shaw, R., Goyal, M., van Embden, J., 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J. Clin. Microbiol. 35, 907–914.

Kapur, V., Whittam, T.S., Musser, J.M., 1994. Is *Mycobacterium tuberculosis* 15,000 years old? J. Infect. Dis. 170, 1348–1349.

Kovalev, S.Y., Kamaev, E.Y., Kravchenko, M.A., Kurepina, N.E., Skomiakov, S.N., 2005. Genetic analysis of *Mycobacterium tuberculosis* strains isolated in Ural region, Russian federation, by MIRU-VNTR genotyping. Int. J. Tuberc. Lung Dis. 9, 746–752.

Kwok, P.Y., 2001. Methods for genotyping single nucleotide polymorphisms. Annu. Rev. Genomics Hum. Genet. 2, 235–258.

Le Fleche, P., Fabre, M., Denoeud, F., Koeck, J.L., Vergnaud, G., 2002. High resolution, on-line identification of strains from the *Mycobacterium tuberculosis* complex based on tandem repeat typing. BMC Microbiol. 2, 37.

Mazars, E., Lesjean, S., Banuls, A.L., Gilbert, M., Vincent, V., Gicquel, B., Tibayrenc, M., Locht, C., Supply, P., 2001. High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology. Proc. Natl. Acad. Sci. U.S.A. 98, 1901–1906.

Millet, J., Miyagi-Shiohira, C., Yamane, N., Sola, C., Rastogi, N., 2007. Assessment of mycobacterial interspersed repetitive unit-QUB markers to further discriminate the Beijing genotype in a population-based study of the genetic diversity of *Mycobacterium tuberculosis* clinical isolates from Okinawa, Ryukyu Islands, Japan. J. Clin. Microbiol. 45, 3606–3615.

Mostowy, S., Cousins, D., Brinkman, J., Aranaz, A., Behr, M.A., 2002. Genomic deletions suggest a phylogeny for the *Mycobacterium tuberculosis* complex. J. Infect. Dis. 186, 74–80.

Musser, J.M., Amin, A., Ramaswamy, S., 2000. Negligible genetic diversity of *Mycobacterium tuberculosis* host immune system protein targets: evidence of limited selective pressure. Genetics 155, 7–16.

Nakajima, H., 1993. Tuberculosis: a global emergency. World Health Forum 14, 438.

Nouvel, L.X., Dos Vultos, T., Kassa-Kelembho, E., Rauzier, J., Gicquel, B., 2007. A non-sense mutation in the putative anti-mutator gene ada/alkA of *Mycobacterium tuberculosis* and *M. bovis* isolates suggests convergent evolution. BMC Microbiol. 7, 39.

Ralph, A.P., Anstey, N.M., Kelly, P.M., 2009. Tuberculosis into the 2010s: is the glass half full? Clin. Infect. Dis. 49, 574–583.

Rohani, M., Farnia, P., Nasab, M.N., Moniri, R., Torfeh, M., Amiri, M.M., 2009. Beijing genotype and other predominant *Mycobacterium tuberculosis* spoligotypes observed in Mashhad city, Iran. Indian J. Med. Microbiol. 27, 306–310.

Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.

Schork, N.J., Fallin, D., Lanchbury, J.S., 2000. Single nucleotide polymorphisms and the future of genetic epidemiology. Clin. Genet. 58, 250–264.

Sebban, M., Mokrousov, I., Rastogi, N., Sola, C., 2002. A data-mining approach to spacer oligonucleotide typing of *Mycobacterium tuberculosis*. Bioinformatics 18, 235–243.

Skuce, R.A., McCorry, T.P., McCarroll, J.F., Roring, S.M., Scott, A.N., Brittain, D., Hughes, S.L., Hewinson, R.G., Neill, S.D., 2002. Discrimination of *Mycobacterium tuberculosis* complex bacteria using novel VNTR-PCR targets. Microbiology 148, 519–528.

Sorek, R., Kunin, V., Hugenholtz, P., 2008. CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. Nat. Rev. Microbiol. 6, 181–186.

Sreevatsan, S., Pan, X., Stockbauer, K.E., Connell, N.D., Kreiswirth, B.N., Whittam, T.S., Musser, J.M., 1997. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. Proc. Natl. Acad. Sci. U.S.A. 94, 9869–9874.

Stavrum, R., Mphahlele, M., Ovreas, K., Muthivhi, T., Fourie, P.B., Weyer, K., Grewal, H.M., 2009. High diversity of *Mycobacterium tuberculosis* genotypes in South Africa and preponderance of mixed infections among ST53 isolates. J. Clin. Microbiol. 47, 1848–1856.

Styblo, K., Bumgarner, J., 1991. Tuberculosis can be controlled with existing technologies: evidence. Tuberculosis Surveillance Research Unit of the IUATLD Progress Report 2, 60–72.

Supply, P., Allix, C., Lesjean, S., Cardoso-Oelemann, M., Rusch-Gerdes, S., Willery, E., Savine, E., de Haas, P., van Deutekom, H., Roring, S., Bifani, P., Kurepina, N., Kreiswirth, B., Sola, C., Rastogi, N., Vatin, V., Gutierrez, M.C., Fauville, M., Niemann, S., Skuce, R., Kremer, K., Locht, C., van Soolingen, D., 2006. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 44, 4498–4510.

Supply, P., Lesjean, S., Savine, E., Kremer, K., van Soolingen, D., Locht, C., 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.

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

Supply, P., Warren, R.M., Banuls, A.L., Lesjean, S., Van Der Spuy, G.D., Lewis, L.A., Tibayrenc, M., Van Helden, P.D., Locht, C., 2003. Linkage disequilibrium between minisatellite loci supports clonal evolution of *Mycobacterium tuberculosis* in a high tuberculosis incidence area. Mol. Microbiol. 47, 529–538.

Takashima, T., Iwamoto, T., 2006. New era in molecular epidemiology of tuberculosis in Japan. Kekkaku 81, 693–707.

van Embden, J.D., Cave, M.D., Crawford, J.T., Dale, J.W., Eisenach, K.D., Gicquel, B., Hermans, P., Martin, C., McAdam, R., Shinnick, T.M., et al., 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J. Clin. Microbiol. 31, 406–409.

van Embden, J.D.A., van Gorkom, T., Kremer, K., Jansen, R., van der Zeijst, B.A.M., Schouls, L.M., 2000. Genetic variation and evolutionary origin of the Direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. J. Bacteriol. 182, 2393–2401.

Vitol, I., Driscoll, J., Kreiswirth, B., Kurepina, N., Bennett, K.P., 2006. Identifying *Mycobacterium tuberculosis* complex strain families using spoligotypes. Infect. Genet. Evol. 6, 491–504.

Warren, R.M., Streicher, E.M., Sampson, S.L., Van Der Spuy, G.D., Richardson, M., Nguyen, D., Behr, M.A., Victor, T.C., Van Helden, P.D., 2002. Microevolution of the direct repeat region of *Mycobacterium tuberculosis*: implications for interpretation of spoligotyping data. J. Clin. Microbiol. 40, 4457–4465.

Wirth, T., Hildebrand, F., Allix-Beguec, C., Wolbeling, F., Kubica, T., Kremer, K., van Soolingen, D., Rusch-Gerdes, S., Locht, C., Brisse, S., Meyer, A., Supply, P., Niemann, S., 2008. Origin, spread and demography of the *Mycobacterium tuberculosis* complex. PLoS Pathog. 4, e1000160.

Zhang, J., Abadia, E., Refregier, G., Tafaj, S., Boschiroli, M.L., Guillard, B., Andremont, A., Ruimy, R., Sola, C., 2010. *Mycobacterium tuberculosis* complex CRISPR genotyping: improving efficiency, throughput and discriminative power of 'spoligotyping' with new spacers and a microbead-based hybridization assay. J. Med. Microbiol. 59, 285–294.

Zozio, T., Allix, C., Günal, S., Saribas, S., Alp, A., Durmaz, R., Fauville-Dufaux, M., Rastogi, N., Sola, C., 2005. The previously identified "LAM7" genotype typifies a Turkish-phylogeographically specific clonal complex of *Mycobacterium tuberculosis*, 26th Annual Congress of the European Society for Mycobacteriology June 26th–29th 2005. Abstract Book, O-8 *presenting author, Istanbul, Turkey, 26–29 June 2005, p. 31.

Nakajima et al. BMC Infectious Diseases 2010, 10:118
http://www.biomedcentral.com/1471-2334/10/118

BMC Infectious Diseases

RESEARCH ARTICLE Open Access

Identification of *Mycobacterium tuberculosis* clinical isolates in Bangladesh by a species distinguishable multiplex PCR

Chie Nakajima*†[1], Zeaur Rahim*†[2], Yukari Fukushima[1], Isamu Sugawara[3], Adri GM van der Zanden[4], Aki Tamaru[5] and Yasuhiko Suzuki*[1]

Abstract

Background: Species identification of isolates belonging to the *Mycobacterium tuberculosis* complex (MTC) seems to be important for the appropriate treatment of patients, since *M. bovis* is naturally resistant to a first line anti-tuberculosis (TB) drug, pyrazinamide, while most of the other MTC members are susceptible to this antimicrobial agent. A simple and low-cost differentiation method was needed in higher TB burden countries, such as Bangladesh, where the prevalence of *M. bovis* among people or cattle has not been investigated.

Methods: Genetic regions *cfp32*, RD9 and RD12 were chosen as targets for a species distinguishable multiplex PCR and the system was evaluated with twenty reference strains of mycobacterial species including non-tubercular mycobacteria (NTM). A total of 350 clinical MTC isolates obtained in Bangladesh were then analyzed with this multiplex PCR.

Results: All of the MTC reference strains gave expected banding patterns and no non-specific amplifications were observed in the NTM strains. Out of 350 clinical isolates examined by this method, 347 (99.1%) were positive for all of the *cfp32*, RD9 and RD12 and determined as *M. tuberculosis*. Two isolates lacked *cfp32* PCR product and one lacked RD12, however, those three samples were further examined and identified as *M. tuberculosis* by the sequence analyses of *hsp65* and *gyrB*.

Conclusions: The MTC-discrimination multiplex PCR (MTCD-MPCR) developed in this study showed high specificity and was thought to be very useful as a routine test because of its simplicity. In the current survey, all the 350 MTC isolates obtained from Bangladesh TB patients were determined as *M. tuberculosis* and no other MTC were detected. This result suggested the general TB treatment regimen including pyrazinamide to be the first choice in Bangladesh.

Background

Mycobacterium tuberculosis complex (MTC), including *M. tuberculosis, M. bovis, M. africanum, M. microti, M. pinnipedii, M. caprae, "M. canettii"* and other closely related strains, is a group of causative agents for human and animal tuberculosis (TB) [1,2]. Although the mycobacterial species in MTC are highly similar to each other in DNA level, MTC members differ widely in terms of host tropism, phenotype and pathogenicity [1,3-5]. No further differentiation is usually performed with isolates determined as MTC, however, it seems to be important in some cases for the appropriate management of patients or for an epidemiological purpose. Especially, in the case of *M. bovis* infection, to identify the species in the early stage of diagnosis is essential to avoid inappropriate treatment, since *M. bovis* is naturally resistant to a major anti-TB drug, pyrazinamide [3,4,6], and the standard regimen including this drug has to be altered.

Several rapid identification methods using nucleic acid amplification techniques have been developed and used for the diagnosis of TB [7-9], however, they do not differentiate *M. tuberculosis* from other members of MTC. Recent comparative genomic analyses have provided valuable information on the region of difference (RD) in

* Correspondence: cnakajim@czc.hokudai.ac.jp
[1] Department of Global epidemiology, Hokkaido University Research Center for Zoonosis Control, Kita20-Nishi10, Kita-ku, Sapporo 001-0020, Japan
† Contributed equally
Full list of author information is available at the end of the article

the chromosome of MTC to indicate that specific identification of MTC can be achieved by the detection of these regions [1-3,10]. PCR-based methods targeting RDs can be easily performed in local clinical laboratories with low expense [3].

Bangladesh is one of the highest TB burden countries, where the estimated number of TB incidence in 2007 was 353,000 to be ranked sixth in the world in the WHO report [11]. In this country, a large number of cattle estimated 23 million heads are reared in households especially in rural areas [12]. Some surveys about drug resistant *M. tuberculosis* [13,14] or *M. tuberculosis* epidemiology [15] have been performed, however, no survey about *M. bovis* prevalence among humans as well as cattle has been reported though people are living in a close relationship with cattle [12,16].

In this study, we developed a simple multiplex-PCR system, named MTC-discrimination multiplex PCR (MTCD-MPCR), to distinguish *M. tuberculosis* from other MTC species using RDs, and applied it for clinical isolates derived from TB patients in Bangladesh.

Methods

Bacterial strains and sample collection

For the evaluation of the method, following twenty reference strains, four MTC strains and sixteen nontuberculous mycobacteria (NTM) strains, were obtained from the Research Institute of Tuberculosis, Japan Anti-Tuberculosis Association (Tokyo, Japan) and used: *M. tuberculosis* H37Rv, *M. africanum* KK13-02, *M. microti* ATCC19422, *M. bovis* BCG Tokyo, *M. avium* JATA51-01, *M. intracellulare* JATA52-01, *M. kansasii* KK21-01, *M. xenopi* KK42-02, *M. fortuitum* JATA61-01, *M. lentiflavum* JATA9N-01, *M. simiae* KK23-01, *M. gordonae* JATA33-01, *M. marinum* JATA22-01, *M. asiaticum* KK24-01, *M. scrofulaceum* JATA31-01, *M. szulgai* JATA32-01, *M. nonchromogenicum* JATA45-01, *M. malmoense* JATA47-01, *M. chelonae* JATA62-01 and *M. abscessus* JATA63-01.

Clinical samples were collected in hospitals in Dhaka, located in an urban area, and Matlab and Sylhet, located in rural areas. A total of 350 isolates, 300 from Dhaka, 41 from Matlab and 9 from Sylhet, were examined (Additional file 1). Among them, 327 isolates were derived from sputa, 22 were from lymph nodes and remaining 1 was from a surgical injury.

Cultivation and biochemical characterization of isolates and DNA extraction

Sputa and other samples were collected from TB suspected patients and decontaminated following the Petropff's method [17]. After a centrifugation at 500 rpm at 4°C, the supernatant was discarded and one loop-full (5 mm diameter of the loop) decontaminated pellet was inoculated onto 2 Lowenstein-Jensen (L-J) slants each. Inoculated L-J slants were incubated at 37°C. Each L-J slant was examined once a week for contamination as well as for growth until 8 weeks. Typical mycobacterium-like colonies were tested for sensitivity to p-nitrobenzoic (PNB) acid. PNB sensitive strains were considered to be *M. tuberculosis* complex. DNA was extracted from those colonies by heating at 95°C for 5 min followed by chloroform extraction and ethanol precipitation [18].

MTC-discrimination multiplex PCR (MTCD-MPCR)

Primer pairs for *cfp32* (Rv0577), RD9 (Rv2073c) and RD12 (Rv3120) designed by Huard *et al* [3] were slightly modified and used as a primer mixture for three simultaneous PCRs in one tube (Table 1). The reaction mixture contained 1 mM dNTPs (0.25 mM each), 0.5 M betaine, 750 nM each of *cfp32* primers (Rv0577F and Rv0577R), 250 nM each of RD primers (Rv2073cF, Rv2073cR, Rv3120F and Rv3120R (390-369)), 1 μL of DNA sample and 1 U of GoTaq DNA Polymerase (Promega Corp., WI, U.S.A.) in 20 μL of Green GoTaq Reaction Buffer. PCR reaction was initiated by denaturation for 1 min at 96°C, followed by 35 cycles of 10 s at 96°C, 20 s at 60°C and 1 min at 72°C with final extension for 5 min at 72°C in a thermalcycler (iCycler, Bio Rad Laboratories Inc., CA, U.S.A.). Reaction mixtures with *M. tuberculosis* DNA and without template DNA were also run simultaneously with samples every time as a positive control and a negative control to evaluate the MTCD-MPCR system. The products were electrophoresed in 2.0% agarose gel in TAE buffer, and stained with ethidium bromide.

The sensitivity of the method was determined using serially diluted purified genomic-DNA solutions, ten-fold dilution from 5 ng/μL to 50 fg/μL, extracted from *M. tuberculosis* H37Rv and *M. bovis* BCG Tokyo. For the specificity study, the concentration of the DNA solution from each reference strain was adjusted to 5 ng/μL and used.

A detection study from sputum samples was performed with *M. tuberculosis* H37Rv spiked sputa. Serially diluted bacteria were spiked into healthy volunteer's sputum samples to final concentrations ranging from 1.5 × 10/mL to 1.5 × 10^6/mL. The sputum was processed by a conventional method and DNA was extracted by bead-beating. Briefly, twice the volume of N-acetyl-L-cysteine and NaOH (NALC-NaOH) was added to the sputum, mixed well and incubated for 15 minutes. The sample was diluted to five times its original volume with PBS and centrifuged for 20 min at 3000 rpm. The sediment was dissolved in 0.75 mL of Tris-EDTA buffer, added 0.5 g of glass beads (0.15 - 0.25 mm, Fuji Chemical Industry Co., Ltd., Japan) and 0.5 mL of chloroform and then shaken with a bead-beater (FastPrep FP100A, MP Biomedicals, U.S.A.) for 5 min at 5500 rpm. Tubes were centrifuged at

5000 rpm for 5 min and 0.5 mL of PCI (phenol-chloroform-isoamyl alcohol, 25:24:1) was added to the supernatant, mixed and centrifuged. DNA was extracted from the PCI treated sample by isopropyl alcohol precipitation and the precipitant was dissolved in 10 μL of Tris-EDTA buffer. Final solutions were subjected to the MTCD-MPCR.

Other PCRs and Sequence analyses

Additional PCRs and sequencings were performed with primers listed in Table 1. Reaction solution components in those PCRs were as follows: 1 mM dNTPs (0.25 mM each), 0.5 M betaine, 500 nM each of forward and reverse primers, 1 μL of DNA sample and 1 U of GoTaq DNA Polymerase (Promega Corp.) in 20 μL of Green GoTaq Reaction Buffer. PCR reaction other than *rrs* was initiated by denaturation for 1 min at 96°C, followed by 30 cycles of 10 s at 96°C, 10 s at 55°C and 30 sec at 72°C with final extension for 5 min at 72°C in a thermalcycler (iCycler, Bio Rad Laboratories Inc.). In *rrs* PCR, the period of extension at 72°C in the cycle was 90 sec, whereas other conditions were same with other PCR procedure. The products were electrophoresed in 1.5% agarose gel and stained with ethidium bromide. Sequencing of PCR product was performed according to manufacturer's protocol with ABI PRISM 3130*xl* Genetic Analyzer (Life Technologies Corp., CA, U.S.A.) using BigDye Terminator v3.1 Cycle Sequencing Kit (Life Technologies Corp.). *rpoB* and *rrs* sequences read by each forward primer (Table 1. *b*) were compared with the sequence of *M. tuberculosis* H37Rv and the samples met following criteria were identified as MTC: *rpoB*, more than 98% match in minimum 150-base length; *rrs*, 100% match in minimum 300 bases. About the samples determined as non-MTC by these criteria, another PCR and sequencing were done with more

Table 1: Used primers for the MTCD-MPCR and additional PCRs and sequencings.

Target locus	Primer name	Primer sequence	Location[a]	Size (bp)	Ref. No.
MTCD-MPCR					
cfp32	Rv0577F	5' ATGCCCAAGAGAAGCGAATACAGGCAA	671166-192	786	[3]
	Rv0577R	5' CTATTGCTGCGGTGCGGGCTTCAA	671951-928		
RD9	Rv2073cF	5' TCGCCGCTGCCAGATGAGTC	2330579-598	600	[3]
	Rv2073cR	5' TTTGGGAGCCGCCGGTGGTGATGA	2331173-150		
RD12	Rv3120F	5' GTCGGCGATAGACCATGAGTCCGTCTCCAT	3485558-587	404	[3]
	Rv3120R (390-369)	5' GCGAAAAGTGGGCGGATGCCAG	3485961-940		
Additional PCRs or sequencings					
cfp32	3'cfp32F	5' CGAATCATTGGCACGTCTACTTTG	671770-793	372	[2]
	3'cfp32R	5' GTGGCACCGGCGGCACCGCACACCT	672141-117		
RD12	Rv3120-F (90-110)	5' GGTATTTGCGCCCATATCCTG	3485661-681	411	this study
	Rv3120-R (500-481)	5' CCTGGCTTCAAGCACCATTC	3486071-052		
rpoB	*rpoB*-F[b]	5' CAGGACGTGGAGGCGATCAC	761007-026	250	[18]
	rpoB-R	5' CAGGGGTTTCGATCGGGCAC	761256-237		
rpoB[c]	*rpoB*-S-F[b]	5' GCGTACGGTCGGCGAGCTGATCC	760922-944	418	this study
	rpoB-S-R	5' GCGGTACGGCGTTTCGATGAACC	761339-317		
rrs	Bact-*rrs*-F[b]	5' AGAGTTTGATCCTGGCTCAG	1471856-875	1496	this study
	Bact-*rrs*-R	5' TACGGCTACCTTGTTACGAC	1473351-332		
rrs[c]	*rrs*-S-F[b]	5' ATACCTTTGGCTCCCTTTTCC	1471809-829	1607	this study
	rrs-S-R	5' CCCACCAGTTGGGGCGTTTTC	1473415-395		
hsp65	*hsp65*F[b]	5' ACCAACGATGGTGTGTCCAT	528752-771	441	[2]
	*hsp65*R	5' CTTGTCGAACCGCATACCCT	529192-173		
gyrB	*gyrB*F[b]	5' ACATCAACCGCACCAAGAACGC	6027-048	483	this study
	*gyrB*R	5' GTGCCTTACGTGCCGCGATACG	6509-488		

[a] Location on the *M. tuberculosis* H37Rv genome (accession no. NC_000962.2).

[b] These primers were used for PCR and sequencing.

[c] Those primer pairs were used for the samples that showed non-MTC gene sequences.

MTC specific primers (Table 1. *c*) with the same criteria. The samples identified as MTC with those specific primers were determined as mixed-culture samples. Species identifications by *hsp65* or *gyrB* sequences were done according to previous publications [2,19,20].

Ethical Clearance

The original research project was approved by the Research Review Committee and Ethical Review Committee of the International Centre for Diaddroeal Disease Research, Bangladesh (ICDDR, B). Signed informed consent was obtained from each patient and volunteer recruited for the study.

Results and discussion

MTCD-MPCR

Three genetic regions were selected as the targets for the multiplex PCR: *cfp32*, RD9 and RD12. *cfp32* is an MTC-restricted gene and used to confirm isolates to belong to MTC [3,21]. RD9 is the region that can be found in only *M. tuberculosis* and "*M. canettii*", and RD12 is found in all MTC members except *M. bovis*, *M. caprae* and "*M. canettii*" [2,3]. By a trial with several patterns of primer concentrations, the best combination was determined to be 750 nM each of *cfp32* primers and 250 nM each of RD9 and RD12 primers in the multiplex PCR. With this PCR, an isolate possessing all the three regions can be identified as *M. tuberculosis* whereas a strain showing only one amplified band, *cfp32*, will be classified as *M. bovis* or *M. caprae* (Figure 1, Table 2). Other banding patterns are interpreted as described in Table 2. "*M. canettii*" is another MTC that clinical isolates reported so far were naturally pyrazinamide resistant [6,19]. Thus, this PCR system was thought to be useful for the discrimination of MTC species, especially for the screening of naturally pyrazinamide-resistant species with only one PCR reaction per sample.

The sensitivities of the method were determined as 500 fg genomic DNA for M. tuberculosis (H37Rv) and 50 fg for M. bovis (BCG Tokyo), which were assumed to be equivalent to 100 bacilli and 10 bacilli, respectively (data not shown). The specificity of the MTCD-MPCR was confirmed with DNA templates extracted from mycobacterial reference strains. Typical gel electrophoresis results of the MTC and NTM strains are shown in Figure 1. All of the PCR products from the MTC strains gave expected banding patterns in correct sizes (Table 1). On the contrary, no bands were obtained from the NTM samples. These results demonstrated the specificity and applicability of the MTCD-MPCR for the differentiation of M. tuberculosis in clinical isolates.

In the study using bacterium-spiked sputa, the detection limit of *M. tuberculosis* was 1.5×10^4 cells in 1 mL of sputum (data not shown). This bacterial concentration is similar or a little higher than the detection limit of the bacteria by the Zeal-Nelsen staining. This means species discrimination by the MTCD-MPCR is applicable for clinical samples if the sputum is diagnosed as smear positive. The discrepancy of the detection limit between purified DNA and bacterium-spiked sputum seems to depend on the extraction or purification procedure, which is to be improved.

Discrimination of MTC in Bangladesh clinical isolates

A total of 350 clinical isolates obtained in Bangladesh were analyzed to see the prevalence of MTC species other than *M. tuberculosis*. All samples were subjected to *rpoB* or *rrs* sequencing to confirm as MTC [22]. By this sequence analysis, 18 samples were revealed as mixtures of MTC and other bacteria, mainly mycobacterial species. Those mixed samples were also subjected to the MTCD-MPCR study.

Out of the 350 isolates, 347 (99.1%) showed the typical banding pattern of *M. tuberculosis* in the electrophoresis by the MTCD-MPCR (Table 2). Among three remaining isolates, one (0.3%, isolate ATP138) lacked RD12 band showing "*M. canettii*" pattern and two (0.6%, isolate M2000 and S2247) did not have *cfp32* band, presented as irregular in Table 2. Those samples were further analyzed to detect the target region, *cfp32* or RD12, with other sets of primers (Table 1) [2]. With this trial, sample S2247 gave an expected sized *cfp32* band, suggesting a partial deletion or a mutation in the primer-binding site. Other two samples, ATP138 and M2000, still failed to amplify the expected DNA fragment indicating that larger deletion event had occurred in the target genomic area. ATP138, the "*M. canettii*" pattern sample, was subjected to *hsp65* sequencing and confirmed not to be "*M. canettii*" [2,19]. These three isolates were further confirmed to be *M. tuberculosis* by the sequencing of *gyrB* [20]. No isolates out of the 350 lacked RD9 band, concordant with the observations by former researchers [2,10,23], possibly indicating the high stability of this region in *M. tuberculosis*. The lack of *cfp32* in an MTC isolate has been reported by Huard *et al* [2] with a similar incidence (1/125, 0.8%) to the current study. Some of the mixed-culture samples showed correct but faint banding patterns, suggesting an inhibitory effect of contaminated DNA. The result indicated that despite its decreased sensitivity, the MTCD-MPCR was able to detect MTC from mixed-culture samples, which sometimes are observed in primary cultures [8].

All of the 350 MTC isolates obtained from clinical specimens in Bangladesh were *M. tuberculosis* and no other MTC species were detected. This information is very helpful for the management of patients to determine treatment regimens. Patients suffered from MTC in surveyed area in Bangladesh can possibly be subjected to the

standard regimen including pyrazinamide. Although *M. bovis* was not detected from human in current study, there are no precise data about *M. bovis* prevalence among cattle in this country, continuous surveys seem to be needed especially in rural areas where people and cattle inhabit more closely [12]. Since pyrazinamide is suggested to have more adverse side effects than other first-line anti-TB drugs [24], useless administration to patients should be avoided.

The MTCD-MPCR developed in this study is considered to be very useful for the differentiation of MTC because of its simplicity and specificity. A large number of samples can be analyzed by this method in a short period of time. Some other MTC discrimination methods using RDs have been published and showed higher differentiation capability that could distinguish almost all members of MTC. However, the procedures were more time consuming (e.g., multiple PCR reactions were needed) [2,3], result interpretations seemed to be complicated (e.g., sizes of amplified bands should be estimated) [10,23,25] or an expensive equipment should be needed [26]. The necessity of detailed MTC discrimination seems to be low since the majority of human tuberculosis causing agents are *M. tuberculosis*, and in some global areas, *M. bovis* partially contributes to the prevalence [4,23]. The vaccine strain bacillus Calmette-Guerin (BCG), an attenuated *M. bovis*, can be a cause of disseminated mycobacterial infection in immunocompromised individuals, however, the patients can be treated by the same regimen as *M. bovis* without pyrazinamide [27], and if necessary, an additional PCR for the detection of RD1 can distinguish BCG from clinical *M. bovis* strains [3]. Possibility of the detection of other MTC species is almost negligible in routine laboratory diagnoses, except for the cases in some African countries where *M. africanum* can be found in a higher ratio [2]. Samples exhibiting rare or irregular banding patterns by the MTCD-MPCR can be examined afterward using other precise methods, e.g., targeting other RDs, gene sequencing or spoligotyping [2,19,20,28]. Thus, the MTCD-MPCR, a simple MTC discrimination method developed and evaluated in this study seemed to be very useful as a screening tool for clinical isolates to distinguish *M. tuberculosis* from *M. bovis* for the prompt decision of treatment regimen.

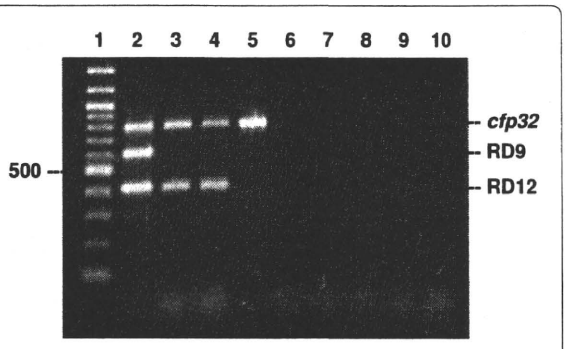

Figure 1 Typical MTCD-MPCR banding patterns of mycobacterial reference strains. PCR products by the MTCD-MPCR were analyzed on 2.0% agarose gel electrophoresis followed by ethidium bromide stain. Lanes: 1, 100-bp ladder; 2, *M. tuberculosis*; 3, *M. africanum*; 4, *M. microti*; 5, *M. bovis*; 6, *M. kansasii*; 7, *M. avium*; 8, *M. intracellulare*; 9, *M. fortuitum*; 10, negative control

Conclusions

In the current study, the MTC-discrimination multiplex PCR (MTCD-MPCR) was developed and applied for a

Table 2: Results of the MTCD-MPCR with Bangladesh clinical isolates.

Species Interpretation	Banding pattern *a*			Number of isolates	%
	cfp32	RD9	RD12		
M. tuberculosis	+	+	+	347	99.1
M. bovis/M. caprae	+	-	-	0	0
"M. canettii"	+	+	-	1 [c]	0.3
Other MTC [b]	+	-	+	0	0
Non MTC	-	-	-	0	0
Irregular	-	+	+	2 [d]	0.6
Total				350	100

[a] Banding patterns are shown as amplification results of the MTCD-MPCR in order of *cfp32*, RD9 and RD12: +, Amplification positive; -, negative.

[b] Other MTC includes following species: *M. africanum, M. microti, M. pinnipedii*, dassie bacillus and oryx bacillus. The latter two are minor strains belonging to MTC [2].

[c] Confirmed to be *M. tuberculosis by hsp65* and *gyrB* sequencing.

[d] Confirmed to be *M. tuberculosis by gyrB* sequencing.

study to see the prevalence of MTC species other than *M. tuberculosis* in clinical isolates in Bangladesh. The method showed high specificity and sensitivity, as 99.1% (347/350) of clinical *M. tuberculosis* isolates were identified by a typical banding pattern. It seemed to be very useful as a routine test method because of its simplicity. All the 350 MTC isolates derived from Bangladesh patients were *M. tuberculosis* and no other MTC was detected. The result suggested that a standard TB treatment regimen including pyrazinamide can be applied to the patients as the first choice in surveyed areas in Bangladesh.

Additional material

Additional file 1 Used isolates in current study. Background information of used isolates including sampling site and drug-susceptibility test results.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

CN, ZR and YS were responsible for planning the study, analyzing the results and drafting the manuscript. YF carried out the molecular genetic studies. AT performed the detection study from sputum samples. ZR, IS, AGMZ and YS collected the study material and coordinated the study. All authors read and approved the manuscript.

Acknowledgements

This work was supported by the International Centre for Diarrhoeal Disease Research, Bangladesh, grants received from Gates-Government of the Peoples Republic of Bangladesh to ZR, Grants-in-Aid for Program of Founding Research Center for Emerging and Reemerging Infectious Diseases from the Ministry of Education, Culture, Sports, Science, and Technology, Japan (MEXT) to YS, and in part by the Global Center of Excellence (COE) Program, "Establishment of International Collaboration Centers for Zoonosis Control" from MEXT to YS and a grant from U.S.-Japan Cooperative Medical Science Programs to IS and YS.

Author Details

[1]Department of Global epidemiology, Hokkaido University Research Center for Zoonosis Control, Kita20-Nishi10, Kita-ku, Sapporo 001-0020, Japan, [2]Tuberculosis laboratory, International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR, B), GPO Box 128, Dhaka 1000, Bangladesh, [3]Research Institute of Tuberculosis, Japan Anti-Tuberculosis Association, 3-1-24 Matsuyama, Kiyose, Tokyo, Japan, [4]Laboratory for Medical Microbiology and Public Health, P.O.Box 377, Burg. Edo Bergsmalaanl, 7512 AD Enschede, The Netherlands and [5]Osaka Prefectural Institute of Public Health, 1-3-69, Nakamichi, Higashinari-ku, Osaka 537-0025, Japan

Received: 26 August 2009 Accepted: 15 May 2010
Published: 15 May 2010

References

1. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS, Samper S, van Soolingen D, Cole ST: **A new evolutionary scenario for the *Mycobacterium tuberculosis* complex.** *Proc Natl Acad Sci USA* 2002, **99**:3684-3689.
2. Huard RC, Fabre M, de Haas P, Lazzarini LC, van Soolingen D, Cousins D, Ho JL: **Novel genetic polymorphisms that further delineate the phylogeny of the *Mycobacterium tuberculosis* complex.** *J Bacteriology* 2006, **188**:4271-4287.
3. Huard RC, de Oliveira Lazzarini LC, Butler WR, van Soolingen D, Ho JL: **PCR-based method to differentiate the subspecies of the *Mycobacterium tuberculosis* Complex on the basis of genomic deletions.** *J Clin Microbiol* 2003, **41**:1637-1650.
4. Smith NH, Gordon SV, de la Rua-Domenech R, Clifton-Hadley RS, Hewinson RG: **Bottlenecks and broomsticks: the molecular evolution of *Mycobacterium bovis*.** *Nat Rev Microbiol* 2006, **4**:670-681.
5. Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, Whittam TS, Musser JM: **Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination.** *Proc Natl Acad Sci USA* 1997, **94**:9869-9874.
6. Somoskovi A, Dormandy J, Parsons LM, Kaswa M, Goh KS, Rastogi N, Salfinger M: **Sequencing of the *pncA* gene in members of the *Mycobacterium tuberculosis* complex has important diagnostic applications: Identification of a species-specific *pncA* mutation in "*Mycobacterium canettii*" and the reliable and rapid predictor of pyrazinamide resistance.** *J Clin Microbiol* 2007, **45**:595-599.
7. Jonas V, Alden MJ, Curry JI, Kamisango K, Knott CA, Lankford R, Wolfe JM, Moore DF: **Detection and identification of *Mycobacterium tuberculosis* directly from sputum sediments by amplification of rRNA.** *J Clin Microbiol* 1993, **31**:2410-2416.
8. Katila ML, Katila P, Erkinjuntti-Pekkanen R: **Accelerated detection and identification of mycobacteria with MGIT 960 and COBAS AMPLICOR systems.** *J Clin Microbiol* 2000, **38**:960-964.
9. Pandey BD, Poudel A, Yoda T, Tamaru A, Oda N, Fukushima Y, Lekhak B, Risal B, Acharya B, Sapkota B, Nakajima C, Taniguchi T, Phetsuksiri B, Suzuki Y: **Development of an in-house loop-mediated isothermal amplification (LAMP) assay for detection of *Mycobacterium tuberculosis* and evaluation in sputum samples of Nepalese patients.** *J Med Microbiol* 2008, **57**:439-443.
10. Parsons LM, Brosch R, Cole ST, Somoskovi A, Loder A, Bretzel G, van Soolingen D, Hale YM, Salfinger M: **Rapid and simple approach for identification of *Mycobacterium tuberculosis* complex isolates by PCR-based genomic deletion analysis.** *J Clin Microbiol* 2002, **40**:2339-2345.
11. WHO: **Epidemiology.** *WHO Report 2009 Global Tuberculosis Control* 2009:6-33.
12. Haque A: **Bangladesh: Social gains from dairy development.** In *Smallholder dairy development: Lessons learned in Asia* Animal Production and Health Commission for Asia and the Pacific, Food and Agriculture Organization of the United Nations, Regional Office for Asia and the Pacific; 2009:8-21.
13. Storla DG, Rahim Z, Islam MA, Plettner S, Begum V, Myrvang B, Bjune G, Rønnild E, Dahle UR, Mannsåker T: **Drug resistance of *Mycobacterium tuberculosis* in the Sunamganj District of Bangladesh.** *Scand J Infect Dis* 2007, **39**:142-145.
14. Zaman K, Rahim Z, Yunus M, Arifeen SE, Baqui AH, Sack DA, Hossain S, Banu S, Islam MA, Ahmed J, Breiman RE, Black RE: **Drug resistance of *Mycobacterium tuberculosis* in selected urban and rural areas in Bangladesh.** *Scand J Infect Dis* 2005, **37**:21-26.
15. Rahim Z, Zaman K, Zanden AG van der, Möllers MJ, van Soolingen D, Raqib R, Zaman K, Begum V, Rigouts L, Portaels F, Rastogi N, Sola C: **Assessment of population structure and major circulating phylogeographical clades of *Mycobacterium tuberculosis* complex in Bangladesh suggests a high prevalence of a specific subclade of ancient *M. tuberculosis* genotypes.** *J Clin Microbiol* 2007, **45**:3791-3794.
16. Rahim Z, Möllers M, te Koppele-Vije A, de Beer J, Zaman K, Matin MA, Kamal M, Raquib R, van Soolingen D, Baqi MA, Heilmann FG, Zanden AG van der: **Characterization of *Mycobacterium africanum* subtype I among cows in a dairy farm in Bangladesh using spoligotyping.** *Southeast Asian J Trop Med Public Health* 2007, **38**:706-713.
17. Petroff SA: **A new and rapid method for isolation and culture of tubercle bacilli directly from the sputum and feces.** *J Exp Med* 1915, **21**:38-42.
18. Suzuki Y, Katsukawa C, Inoue K, Yin Y, Tasaka H, Ueba N, Makino M: **Mutations in *rpoB* gene of rifampicin resistant clinical isolates of *Mycobacterium tuberculosis* in Japan.** *Kansenshogaku Zasshi* 1995, **69**:413-419.
19. Goh KS, Legrand E, Sola C, Rastogi N: **Rapid differentiation of "*Mycobacterium canettii*" from other *Mycobacterium tuberculosis* complex organisms by PCR-restriction analysis of the *hsp65* gene.** *J Clin Microbiol* 2001, **39**:3705-3708.
20. Niemann S, Harmsen D, Rüsch-Gerdes S, Richter E: **Differentiation of clinical *Mycobacterium tuberculosis* complex isolates by *gyrB* DNA sequence polymorphism analysis.** *J Clin Microbiol* 2000, **38**:3231-3234.

21. Huard RC, Chitale S, Leung M, Lazzarini LC, Zhu H, Shashkina E, Laal S, Conde MB, Kritski AL, Belisle JT, Kreiswirth BN, Lapa e Silva JR, Ho JL: **The *Mycobacterium tuberculosis* complex-restricted gene *cfp32* encodes an expressed protein that is detectable in tuberculosis patients and is positively correlated with pulmonary interleukin-10.** *Infect Immun* 2003, **71**:6871-6883.
22. Kim BJ, Lee SH, Lyu MA, Kim SJ, Bai GH, Chae GT, Kim EC, Cha CY, Kook YH: **Identification of mycobacterial species by comparative sequence analysis of the RNA polymerase gene (*rpoB*).** *J Clin Microbiol* 1999, **37**:1714-1720.
23. Das S, Das SC, Verma R: **Occurrence of RD9 region and 500 bp fragment among clinical isolates of *Mycobacterium tuberculosis* and *Mycobacterium bovis*.** *Microbiol Immunol* 2007, **51**:231-234.
24. Yee D, Valiquette C, Pelletier M, Parisien I, Rocher I, Menzies D: **Incidence of serious side effects from first-line antituberculosis drugs among patients treated for active tuberculosis.** *Am J Respir Crit Care Med* 2003, **167**:1472-1477.
25. Warren RM, Gey van Pittius NC, Barnard M, Hesseling A, Engelke E, de Kock M, Gutierrez MC, Chege GK, Victor TC, Hoal EG, van Helden PD: **Differentiation of *Mycobacterium tuberculosis* complex by PCR amplification of genomic regions of difference.** *Int J Tuberc Lung Dis* 2006, **10**:818-822.
26. Pinsky BA, Banaei N: **Multiplex real-time PCR assay for rapid identification of *Mycobacterium tuberculosis* complex members to the species level.** *J Clin Microbiol* 2008, **46**:2241-2246.
27. Rezai MS, Khotaei G, Mamishi S, Kheirkhah M, Parvaneh N: **Disseminated Bacillus Calmette-Guerin infection after BCG vaccination.** *J Trop Pediatr* 2008, **54**:413-416.
28. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, van Embden J: **Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology.** *J Clin Microbiol* 1997, **35**:907-914.

Pre-publication history
The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2334/10/118/prepub

doi: 10.1186/1471-2334-10-118
Cite this article as: Nakajima *et al.*, Identification of Mycobacterium tuberculosis clinical isolates in Bangladesh by a species distinguishable multiplex PCR *BMC Infectious Diseases* 2010, **10**:118

Jpn. J. Infect. Dis., 63, 246-250, 2010

Original Article

A Novel Method for the Purification of DNA by Capturing Nucleic Acid and Magnesium Complexes on Non-Woven Fabric Filters under Alkaline Conditions for the Gene Diagnosis of Tuberculosis by Loop-Mediated Isothermal Amplification (LAMP)

Tadashi Fukasawa, Naozumi Oda, Yasunao Wada[1], Aki Tamaru[2], Yukari Fukushima[3], Chie Nakajima[3], and Yasuhiko Suzuki[3]*

Asahi-Kasei Co., Ltd., Fuji 416-8501;
[1]Clinical Laboratory, The Hospital of Hyogo College of Medicine, Nishinomiya 663-8501;
[2]Department of Infectious Diseases, Osaka Prefectural Institute of Public Health, Osaka 537-0025; and
[3]Department of Global Epidemiology, Hokkaido University Research Center for Zoonosis Control, Sapporo 001-0020, Japan

(Received March 16, 2010. Accepted May 24, 2010)

SUMMARY: A novel method for purifying DNA from clinical samples based on the complex formation of DNA and magnesium ion (Mg^{2+}) was developed for the detection of *Mycobacterium tuberculosis*. The formation of a DNA-Mg^{2+} complex under alkaline conditions was observed by analyzing electrophoretic mobility reduction of DNA on agarose gel. The DNA-Mg^{2+} complex increases the efficacy of DNA recovery from the sample solution on polyethylene terephthalate non-woven fabric (PNWF) filters. Among the various divalent metal cations, only Mg^{2+} was associated with this effect. The applicability of DNA recovered on the PNWF filter was examined for the gene amplification method; loop-mediated isothermal amplification (LAMP). DNA on the PNWF filter could be used for the amplification of specific DNA fragments without elution from the filter. Using this method, DNA was directly purified from *M. tuberculosis* spiked sputum and examined by LAMP assay, showing a high sensitivity in comparison to the commercially available DNA extraction kit. These results indicated that the method developed in this study is useful for rapid gene diagnosis of tuberculosis.

INTRODUCTION

The detection of pathogens in clinical specimens is very important for treating infectious diseases. Most bacterial pathogens grow rapidly on culture media, and a laboratory diagnosis can be made within 1 to 3 days. On the contrary, the slow growth rate of *Mycobacterium tuberculosis*, taking several weeks to form visible colonies, is a well-known problem for clinical laboratories in its identification. To deal with the urgent need for point-of-care testing for tuberculosis in hospitals and primary care facilities, gene diagnosis appears to be crucial.

Usually, gene diagnostic methods consist of two processes; the purification of DNA and the detection of specific DNA fragments by gene amplification technique. For successful gene amplification, DNA purification is the most important step, as contaminated proteins or others may disturb the amplification of the target DNA fragment. Although a number of gene amplification techniques have been successfully used for gene diagnosis, the purification of nucleic acid from specimens such as sputum is still laborious and time-consuming. Therefore, the development of a simple and easy DNA purification system is required to allow laboratory diagnosis of tuberculosis with ease. In general, the method developed by Boom et al. utilizing the binding ability of DNA to silica or glass particles in the presence of chaotropic agents (NaI or $NaClO_4$) has been used to purify DNA from specimens (1,2). Although this method is very useful, it requires a complicated process of eluting DNA from the silica or glass particles.

Over the last two decades, there has been a great deal of interest in investigating the binding of divalent metal cations to DNA because of the need for concentrations of small amounts of DNA for diagnostic use and a purification of a large amount of plasmids for therapeutic use. Metal cations cause DNA condensation with polyamine (spermine, spermidine) or alcohol (3-7). Magnesium ion (Mg^{2+}) is known to interact with DNA and to cause morphological changes (8,9). Because Mg^{2+} in aqueous solution could easily precipitate as magnesium hydroxide under alkaline conditions, the interaction between Mg^{2+} and DNA under high pH conditions has never been investigated.

In this paper, the formation of a DNA and Mg^{2+} complex under alkaline conditions was investigated. The recovery of DNA-Mg^{2+} complex from the sample solution on polyethylene terephthalate non-woven fabric (PNWF) filters was examined as a possible rapid, simple, and low cost method to purify and concentrate DNA obtained from decontaminated clinical specimens of tuberculosis patients for an isothermal gene diagnos-

*Corresponding author: Mailing address: Department of Global Epidemiology, Hokkaido University Research Center for Zoonosis Control, Kita 20-Nishi 10, Kita-ku, Sapporo 001-0020, Japan. Tel: +81-11-706-9503, Fax: +81-11-706-7310, E-mail: suzuki@czc.hokudai.ac.jp

tic method, loop-mediated isothermal amplification (LAMP).

MATERIALS AND METHODS

Filter: PNWF filter Product No. A040C01 (Asahi-Kasei, Shizuoka, Japan) with a thickness of 235 μm and an average pore size of 2.78 μm, was used for DNA recovery. The PNWF filter was cut into 5-mm diameter circles, attached to a QSP Yellow Pipette Tip (Porex Bio, Fairburn, Ga., USA) and termed as a PNWF filter tip. This was connected to a plastic syringe (Terumo, Tokyo, Japan) via vinyl tubing for use as in Fig. 1.

DNA: Purified *Neisseria gonorrhoeae* genomic DNA was purchased from the American Type Culture Collection (no. 700825) and used as a template for the LAMP reaction. λ Phage DNA was purchased from Sigma-Aldrich (St. Louis, Mo., USA).

Divalent metal cations: Metal(II) chloride compounds, $MgCl_2$, $CaCl_2$, $SrCl_2$, $CoCl_2$, $NiCl_2$, $CuCl_2$, $FeCl_2$, and $ZnCl_2$ were purchased from Wako Pure Chemicals (Osaka, Japan). $BaCl_2$ was purchased from Sigma-Aldrich. All metal chloride compounds that were used for the experiments were prepared as 100 mM aqueous stock solutions and then used.

LAMP: The LAMP reaction mixture consists of 2.5 μM (F3, B3 primer), 20 μM (FIP, BIP primer), 30 μM (F-Loop, B-Loop primer), 1.4 mM dNTPs, 0.8 M trimethylglycine, 20 mM Tris-HCl (pH 8.8), 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Tween20, 6.4 units *Bst* DNA polymerase large fragment (New England Biolabs, Beverly, Mass., USA), and 8 mM $MgSO_4$. The volume of the reaction mix was adjusted to 50 μL. The primer sets for LAMP detection of *N. gonorrhoeae* and *M. tuberculosis* were designed for targeting the ORF1 (Genbank accession no. S86113) of *N. gonorrhoeae* and the 16S ribosomal RNA gene of *M. tuberculosis*. The primer sequences are presented in Table 1. All primers used in this study were synthesized by Nihon Gene Research Laboratories (Miyagi, Japan). The LAMP reactions were conducted at 64°C for 1 h, and DNA amplification was monitored based on turbidity at 650 nm with a real time turbid meter LA-200 (Eiken Chemical, Tokyo, Japan).

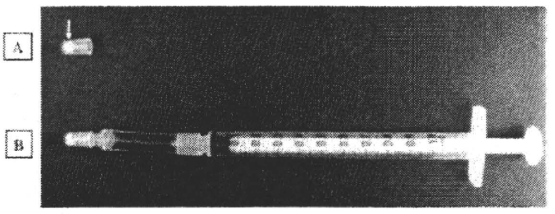

Fig. 1. A device for the entrapment of DNA-Mg^{2+} complex. The PNWF filter was cut into 5-mm diameter circles, attached to a trancated QSP Yellow Pipette Tip and termed as a PNWF-filter tip (A). This was connected to a 1-mL plastic syringe via vinyl tubings (B). PNWF-filter tip was removed from the device after entrapping of DNA and transferred into a tube with LAMP reaction mixture. White arrow indicated a PNWF filter on a tip.

Interaction between DNA and divalent metal cations: A preparation of 0.8% agarose gel was prepared by dissolving 0.8 g agarose (Invitrogen Japan, Tokyo, Japan) into TA (40 mM Tris-acetate, pH 7.5) buffer or 0.05 N NaOH solution. For electrophoresis under alkaline conditions, 600 ng λ phage DNA containing divalent metal cations were applied to 0.05 N NaOH agarose gels with an alkaline loading buffer (13% sucrose, 0.025% bromocresol green, 0.04% xylene cyanol, 0.05 N NaOH) at various concentrations of $MgCl_2$ or 1 mM metal [II] chlorides, and then electrophoresed at 50 V for 3 h in 0.05 N NaOH. After electrophoresis, the gels were neutralized by immersion in neutralization buffer (1 M Tris-HCl pH 8.0, 1.5 M NaCl) for 30 min. The DNA mobility shifts on the agarose gels were assessed after staining for 30 min with 1 μg/ml ethidium bromide in a neutralization buffer by Bio Image Gel Print2000i/VGA (Shimadsu Biotech, Kyoto, Japan). For electrophoresis under normal pH conditions, 50 ng λ phage DNA containing divalent metal cations with loading buffer (13% sucrose, 0.025% bromocresol green, 0.04% xylene cyanol) were applied to agarose gels prepared in TA buffer. The mobility shifts of the DNA on the agarose gel were assessed after staining with 1 μg/ml ethidium bromide in distilled water.

LAMP reaction with recovered DNA on a PNWF filter: One milliliter saline solution containing *N. gonorrhoeae* genome DNA (20 pg) with any one of 0.5 N NaOH, 0.3% bovine serum albumin (BSA), 1 mM

Table 1. LAMP primers for *N. gonorrhoeae* and *M. tuberculosis*

Target	Name	Sequence*	Reference
N. gonorrhoeae	NG-FIP	cgcccaaacagtttcacaacctattttcaggatgtggcgg	This study
	NG-BIP	cgatttctcccattgggctccttctacgatgacatcgc	
	NG-F3	cttaattcttctagtaacaaacc	
	NG-B3	gggaatagttggatcattcgg	
	NG-FL	ccagcttgatgaaagccc	
	NG-BL	ggcttgcgaaagtttccg	
M. tuberculosis	TB-FIP	cacccacgtgttactcatgcaagtcgaacggaaaggtct	Pandy et al. (11)
	TB-BIP	tcgggataagcctggaccacaagacatgcatcccgt	
	TB-F3	gttcgccactcgagtatctccg	
	TB-B3	gaaactgggtctaataccgg	
	TB-FL	ctggctcaggacgaacg	
	TB-BL	gctcatcccacaccgc	

*Sequences of oligonucleotides are from 5′ to 3′.

$MgCl_2$, or 1 mM EDTA in a volume of 1 ml was passed through a PNWF filter using a plastic syringe. As a model of protein-containing samples, 0.3% BSA (Sigma-Aldrich) was added. The DNA recovered on the PNWF filter was washed by passing with 1 mL TBS buffer (10 mM Tris-HCl pH 8.0, 150 mM NaCl). Then, the PNWF-filter tip was transferred into the Loopamp reaction tube (Eiken Chemical) with 50 μL LAMP reaction mix. The LAMP reactions were carried out as described above.

Detection of *M. tuberculosis* in spiked sputum: Mycobacteria-negative sputum from a volunteer was aliquoted in 0.2 mL and spiked with 10 μL of 5 × serially diluted *M. tuberculosis* H37Rv suspension from 4,000,000 to 256 bacilli/mL in phosphate buffer. Each sample was then decontaminated by *N*-acetyl-L-cysteine (NALC)-NaOH procedure recommended by Centers for Disease Control and Prevention, USA, to obtain precipitate of tubercle bacilli. The precipitates were suspended in 200 μL of saline and divided into 2 PCR tubes and used for DNA preparation. A part (100 μL) was treated with 3 cycles of heating at 95°C for 5 min and freezing at −80°C for 5 min. Fifty microliters of freeze-thaw samples were then mixed with 100 μL of distilled water, 10 μL of 10 mM $MgCl_2$ and 40 μL of 5 N NaOH. Samples with a final volume of 200 μL were passed through the PNWF filter attached to a plastic syringe. And the PNWF filters were washed by passing through 1 mL Tris-buffered saline using a plastic syringe. The PNWF filters were then directly transferred into the LAMP reaction tube with 50 μL LAMP reagent mixture. LAMP reactions were carried out as described above. Alternatively, DNA extract was obtained using DNA extraction module in Amplicore *Mycobacteria*™ (Roche Diagnostics, Tokyo, Japan) according to the manufacturer's manual. Briefly, 100 μL of suspension from decontaminated sputum was mixed with 500 μL of washing solution in a micro centrifuge tube, vortexed and centrifuged at 12,500 × g for 10 min. The precipitate was then suspended in 50 μL specimen lysis reagent, vortexed and incubated at 60°C for 45 min. Then, 50 μL specimen neutralization reagent was added. Five microliters of resulting DNA extract was then subjected to LAMP reaction in a total volume of 50 μL. Results were expressed as positive (+) or negative (−) according to the turbidity of more than 0.1 at 650 nm after 60 min incubation at 64°C.

RESULTS AND DISCUSSION

To establish a DNA purification method suitable for gene amplification by LAMP, the effect of divalent metal cations under alkaline conditions for retaining DNA on PNWF filter was investigated.

Effect of metal cations on DNA mobility: The effects of various divalent metal cations on the mobility of λ phage DNA under alkaline conditions were investigated (Fig. 2). Cu and Zn cations showed no effect on the mobility of λ phage DNA under alkaline conditions. A small amount of λ phage DNA was observed to stack at the origin of agarose gel in the presence of Ca, Sr, and Ba cations. Smear bands observed with Co^{2+} and Ni^{2+} indicated the minor effect of these cations. λ Phage DNA band was not observed but smearing DNA bands were observed due to DNA fragmentation with Fe^{2+}. This degradation might be caused by oxygen radicals elicited during the conversion of Fe^{2+} to Fe^{3+} (10). These results suggested that λ phage DNA was prevented from migrating into the gels by the addition of Mg^{2+} under alkaline conditions, while other divalent metal cations examined (alkaline earth metals, transition metals) had a low or zero effect on the migration of DNA in agarose. The DNA mobility reduction data suggested that only Mg^{2+} could form an efficient complex with DNA under alkaline conditions.

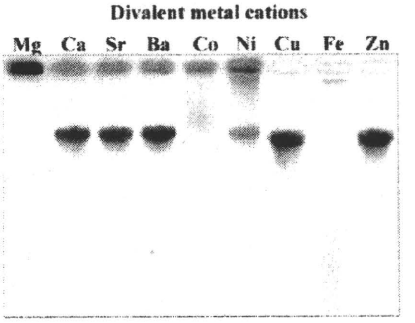

Fig. 2. Electrophoretic mobility shift of λ phage DNA (48 kbp) with divalent metal cations. 600 ng λDNA samples with 1 mM divalent metal cations, Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Fe^{2+}, and Zn^{2+} were electrophoresed in agarose gel containing 50 mM NaOH to reveal electrophoretic mobility shift.

Table 2. Effect of Metal (II) cations upon the retention of DNA on non-woven fabric filter

Metal (II) cation	Threshold time[1)] (min)
Mg^{2+}	19.9 ± 1.3
Ca^{2+}	40.8 ± 18.2
Sr^{2+}	33.0 ± 23.5
Ba^{2+}	51.6 ± 14.5
Co^{2+}	NA[2)]
Ni^{2+}	48.4 ± 10.0
Cu^{2+}	NA
Fe^{2+}	53.1 ± 11.9
Zn^{2+}	39.9 ± 20.6

1): The time needed for the turbidity of each sample to exceed 0.1 at 650 nm with a real time turbid meter.
2): NA, amplification was not observed after 60 min.

Effect of various divalent metal cations on DNA retention on the PNWF filter: Next, the effect of various divalent metal cations on the DNA retention efficiency on the PNWF filter was examined. After adding various divalent metal cations to *N. gonorrhoeae* DNA samples containing BSA, the DNA was adsorbed on the PNWF filter. To evaluate the efficiency of DNA retention, the LAMP reaction for *N. gonorrhoeae* ORF1 was performed with the PNWF filter. The smaller LAMP reaction time means higher retention efficiency. Among the various divalent metal cations examined, Mg^{2+} resulted in the highest recovery of DNA (Table 2).

Effect of Mg^{2+} on DNA recovery on PNWF filter: The interaction between DNA and divalent metal cations under alkaline conditions was investigated using agarose gel electrophoresis. The mobility of λ phage