CVI
Journals.ASM.org

Gene Expression Profile and Immunological Evaluation of Unique Hypothetical Unknown Proteins of *Mycobacterium leprae* by Using Quantitative Real-Time PCR

Hee Jin Kim, Kalyani Prithiviraj, Nathan Groathouse, Patrick J. Brennan, John S. Spencer

Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, Colorado, USA

The cell-mediated immunity (CMI)-based *in vitro* gamma interferon release assay (IGRA) of *Mycobacterium leprae*-specific antigens has potential as a promising diagnostic means to detect those individuals in the early stages of *M. leprae* infection. Diagnosis of leprosy is a major obstacle toward ultimate disease control and has been compromised in the past by the lack of specific markers. Comparative bioinformatic analysis among mycobacterial genomes identified potential *M. leprae*-specific proteins called "hypothetical unknowns." Due to massive gene decay and the prevalence of pseudogenes, it is unclear whether any of these proteins are expressed or are immunologically relevant. In this study, we performed cDNA-based quantitative real-time PCR to investigate the expression status of 131 putative open reading frames (ORFs) encoding hypothetical unknowns. Twenty-six of the *M. leprae*-specific antigen candidates showed significant levels of gene expression compared to that of ESAT-6 (ML0049), which is an important T cell antigen of low abundance in *M. leprae*. Fifteen of 26 selected antigen candidates were expressed and purified in *Escherichia coli*. The seroreactivity to these proteins of pooled sera from lepromatous leprosy patients and cavitary tuberculosis patients revealed that 9 of 15 recombinant hypothetical unknowns elicited *M. leprae*-specific immune responses. These nine proteins may be good diagnostic reagents to improve both the sensitivity and specificity of detection of individuals with asymptomatic leprosy.

The diagnosis of leprosy is usually based solely on clinical symptoms, requiring the presence of a neurological deficit and skin lesions (1; http://www.who.int/lep/diagnosis/en/index.html), but due to low sensitivity, physical diagnosis is applicable only to patients with actual disease. More than 70% of infected patients are negative for acid-fast bacilli (AFB) and do not present analgesic skin lesions, especially paucibacillary/tuberculoid (PB/TT) leprosy patients (1). Since the presence of skin lesions in these patients is variable, their clinical symptoms are not sufficient to specifically diagnose leprosy (1). These problems are accentuated in the case of diagnosis of individuals with subclinical *Mycobacterium leprae* infection, including household contacts (HHCs) of leprosy patients, regarded as the primary source of ongoing leprosy prevalence (1–5).

Several *M. leprae* antigens have been identified and evaluated for their diagnostic potential by serological or cell-mediated immunity (CMI)-based tests (2, 4–24). Serological assay with a single *M. leprae*-specific antigen, phenolic glycolipid I (PGL-I), successfully detects circulating antibodies in sera of multibacillary/lepromatous leprosy (MB/LL) patients. However, this test fails to detect the majority of PB/TT patients and HHCs, though these individuals present strong CMI responses to mycobacterial antigens (3, 13). Both *in vitro* gamma interferon release assays (IGRAs) and a simple delayed-type hypersensitivity skin test have been developed to detect individuals in the early stages of leprosy, using highly antigenic *M. leprae* fractions and the major individual immunogenic proteins (4, 8, 14, 22, 25). However, an obstacle in the application of IGRAs to the major *M. leprae* protein antigens is that most of these antigens share appreciable homology with orthologues in *Mycobacterium* spp. in general, resulting in undesirable cross-reactivity in individuals such as those vaccinated with *M. bovis* BCG or exposed to *M. tuberculosis* or nontuberculous mycobacteria.

Comparative genomic analyses of *M. leprae* and other mycobacteria have identified up to 142 hypothetical unknown open reading frames (ORFs) coding for *M. leprae*-specific proteins (hypothetical unknowns) (26, 27). Recently, either recombinant proteins or synthetic peptides originating from these ORFs and containing T cell epitopes restricted via major HLA-DR alleles have been studied as *M. leprae*-specific antigens (6, 7, 14–17, 22), and some have been shown by IGRAs to differentiate individuals infected with *M. leprae* from healthy controls in areas of endemicity (16, 17). However, the levels of gamma interferon (IFN-γ) secretion in response to these antigens, particularly their peptides, were often too low to distinguish all individuals exposed to *M. leprae* from healthy volunteers in regions of endemicity (16, 22). The low sensitivity of current IGRAs raises the question of whether any of the hypothetical unknowns are expressed or are immunologically relevant, especially considering that about 50% of *M. leprae* genes encoding functional proteins in other mycobacteria are deleted or are pseudogenes (26; http://genolist.pasteur.fr/Leproma/; http://genolist.pasteur.fr/TubercuList/).

The aim of this study was to identify *M. leprae*-specific proteins that can be expressed in the *M. leprae* proteome and recognized by the host immune system, to eventually be used as diagnostic re-

Received 3 July 2012 Returned for modification 25 July 2012
Accepted 29 November 2012
Published ahead of print 12 December 2012
Address correspondence to Hee Jin Kim, Hee.Kim@colostate.edu.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/CVI.00419-12.

doi:10.1128/CVI.00419-12

agents to differentiate individuals with asymptomatic *M. leprae* infection, as well as PB/TT patients, from healthy individuals in regions where leprosy is endemic. In order to achieve this goal, we performed cDNA-based quantitative real-time PCR (qRT-PCR) to investigate the expression status of 131 *M. leprae*-specific ORFs (hypothetical unknowns) and selected 26 promising antigen candidates which showed relatively high gene expression levels for recombinant protein production. Subsequent serological analysis using sera from LL patients and cavitary tuberculosis (TB) patients evaluated the immunological potential of these new recombinant antigens.

MATERIALS AND METHODS

Isolation of *M. leprae* RNA. *M. leprae* Thai-53 was isolated from the livers and spleens of experimentally infected armadillos (provided by R. W. Truman, National Hansen's Disease Laboratories) as described by Shepard et al. (21). Bacteria were suspended in a vial containing 1 ml of TRIzol (Invitrogen Life Technologies, Carlsbad, CA) and lysing matrix B (MP Biomedical LLC, Solon, OH) and were mechanically lysed using a Fast Prep-24 instrument (MP Biomedical LLC, Solon, OH) (28). The resulting homogenate was added to 200 μl of chloroform-isoamyl alcohol (24:1 [vol/vol]), mixed, and centrifuged at 27,000 × *g* for 20 min. Nucleic acids in the aqueous phase were precipitated by adding 100 μl of 3 M sodium acetate (pH 5.2) and 500 μl of isopropanol, followed by incubation at −20°C for 1 h. Total RNA was recovered by centrifugation at 27,000 × *g* for 30 min at 4°C. A Turbo DNA-free kit (Ambion, Austin, TX) was used to remove the DNA contaminants in the RNA solution prior to cDNA synthesis, following the manufacturer's instructions.

Primer design for qRT-PCR analysis. DNA sequences of all 131 hypothetical unknown ORFs in functional class VI (26; http://www.pasteur.fr/recherche/unites/Lgmb/NATURE_DATA/ML_gene_list) were obtained from the *M. leprae* genome database, Leproma (http://genolist.pasteur.fr/Leproma/). OLIGO6 primer analysis software (Molecular Biology Insights Inc., Cascade, CO) was used to design specific primers for each target gene among the hypothetical unknown ORFs. The ML2244, ML2249, ML2567, ML2151, ML0567, and ML0678 genes were excluded; the genes encoding ML2249 and ML2151 were too small for design of proper primers, and the gene expression levels of ML0567, ML0678, and ML2567 had already been studied at the time that this work was initiated (14, 29). These three unknowns appeared to be transcribed significantly in *M. leprae* strains isolated from infected mice or lepromatous patients (14, 29). In order to enhance the efficiency of qRT-PCR, the primers for each target gene were designed to produce a PCR product of 200 to 400 bp. The specificity of each primer set for the template was analyzed by comparison with the genomes of *M. tuberculosis* (http://genolist.pasteur.fr/TubercuList/), *M. avium*, *M. bovis* BCG, and *M. smegmatis* (J. Craig Venter Institute [JCVI] Microbial Database [http://cmr.jcvi.org/cgi-bin/CMR/CmrHomePage.cgi]) through BLAST searches (see Table S1 in the supplemental material).

Synthesis of cDNA and qRT-PCR assays. Total RNA transcripts of *M. leprae* Thai-53 were converted to cDNA by use of a SuperScript III first-strand synthesis kit (Invitrogen Life Technologies, Carlsbad, CA) with random hexamers according to the manufacturer's instructions. All PCR mixtures had a final volume of 25 μl and were set up in triplicate in 96-well optical-grade PCR plates (Bio-Rad Laboratories, Hercules, CA). The primer sets for *ml0380*, *ml2038*, and *east-6*, which were previously shown to be expressed in *M. leprae* (22, 29), were used to optimize the PCR conditions. An initial DNA denaturation step at 95°C for 5 min was followed by 30 cycles of denaturation at 95°C for 30 s, primer annealing at 64°C for 20 s, and primer extension at 72°C for 45 s, with a final extension step at 72°C for 5 min. The qRT-PCR assay of each gene was performed with *M. leprae* cDNA and a dilution series of quantified *M. leprae* genomic DNA (10 fg, 50 fg, 100 fg, 500 fg, 1 pg, 10 pg, 100 pg, and 1 ng per μl) for relative quantification of cDNA by using Platinum SYBR green qPCR SuperMix-UDG (Invitrogen Life Technologies, Carlsbad, CA) on an iCycler real-time PCR machine (Bio-Rad Laboratories, Hercules, CA). Each reaction with independent serial dilutions of genomic DNA and cDNA was performed in triplicate. The variance in reaction replicates was <0.05. Average threshold cycle (C_T) values for each reaction and the initial amount of genomic DNA were used to create a standard curve and determine the relative quantities of the target gene in the form of cDNA. Using iCycler iQ software (Bio-Rad Laboratories, Hercules, CA), the efficiency of qRT-PCR for each target was calculated from the slopes of linear standard curves. The correlation coefficiency ($10^{-1/\text{slope}}$) between each target and standard curve was >0.95, along with 85% to 105% PCR efficiency.

Cloning of hypothetical unknown ORFs from *M. leprae*. To express 24 of 26 selected novel antigen candidates in *Escherichia coli*, genes encoding these hypothetical unknowns were PCR amplified from *M. leprae* Thai-53 genomic DNA by using *rTth* DNA polymerase XL (Applied Biosystems, Carlsbad, CA). PCR amplification was performed using primer sets which included NdeI and HindIII sites specific to the upstream and downstream sequences of the open reading frames (see Table S2 in the supplemental material). Each of the target genes was amplified using touchdown PCR. This method had a high initial annealing temperature of 64°C that decreased by an additional 1°C in each of the first 7 cycles, followed by 25 cycles at 58°C. PCR products were directly digested with restriction enzymes and cloned into the expression vector pET29a(+) (Novagen, Madison, WI), which contained the coding sequence for a 6-histidine tag at the C termini of expressed proteins to facilitate the purification of recombinant proteins. The DNA sequences of all recombinant clones were confirmed by automated nucleotide sequencing at the Proteomics and Metabolomics Facilities, Colorado State University.

Purification of recombinant proteins. The plasmids containing the novel antigen candidate genes were introduced into the *E. coli* expression host BL21 Star(DE3)pLysS (Invitrogen Life Technologies, Carlsbad, CA) according to the manufacturer's instructions. The transformants were grown to log phase (optical density at 600 nm of 0.5) at 37°C in Luria-Bertani (LB) broth with 50 μg/ml kanamycin. Expression of recombinant proteins was induced by adding 0.2 to 0.5 mM IPTG (isopropyl-β-D-thiogalactopyranoside). The cells were cultured at 25°C overnight. The cultured cells were harvested by centrifugation at 4°C and frozen at −70°C. The cells were resuspended in lysis buffer (50 mM Tris-HCl [pH 8.0], 300 mM NaCl) containing 10 μg/ml of DNase and 10 μg/ml of RNase, as well as a protease inhibitor cocktail (P8340; Sigma, St. Louis, MO) and 20 μM phenylmethylsulfonyl fluoride (PMSF; Sigma, St. Louis, MO), and were disrupted by intermittent probe sonication with a Soniprep 150 sonicator (Sanyo MSE, London, United Kingdom) for 10 min. The lysates were centrifuged at 5,000 × *g* for 5 min to remove unbroken cells, and the supernatants were centrifuged for 30 min at 27,000 × *g* at 4°C. Supernatants were applied to a Ni-nitrilotriacetic acid (Ni-NTA) agarose column (Qiagen, Valencia, CA) and washed with 20 column volumes of the same buffer, and the recombinant proteins were eluted with 1 column volume of lysis buffer containing stepwise increments of imidazole (22). All recombinant proteins were found in the 50 to 200 mM imidazole fractions. Of the 24 candidates, 15 were expressed and purified in sufficient quantity and purity for subsequent analysis.

Subjects and samples. Serum samples were coded to protect donor identities and collected with informed consent and with permission from the institutional review boards of the relevant countries and institutions involved, as described by Spencer et al. (23). All leprosy patient sera were obtained from newly diagnosed individuals prior to their receiving the multidrug regimen. Based on bacterial index, histological, and clinical observations, leprosy patients were classified according to the Ridley-Jopling classification system (30) and recruited at the Leonard Wood Memorial Center for Leprosy Research, Cebu, Philippines. It was not determined whether the LL patients developed leprosy type 1 or type 2 reactions. Sera from cavitary TB patients were provided by William MacKenzie through a serum bank repository from the Centers for Disease Control, Atlanta, GA (23). All TB patients were smear positive. Sera were

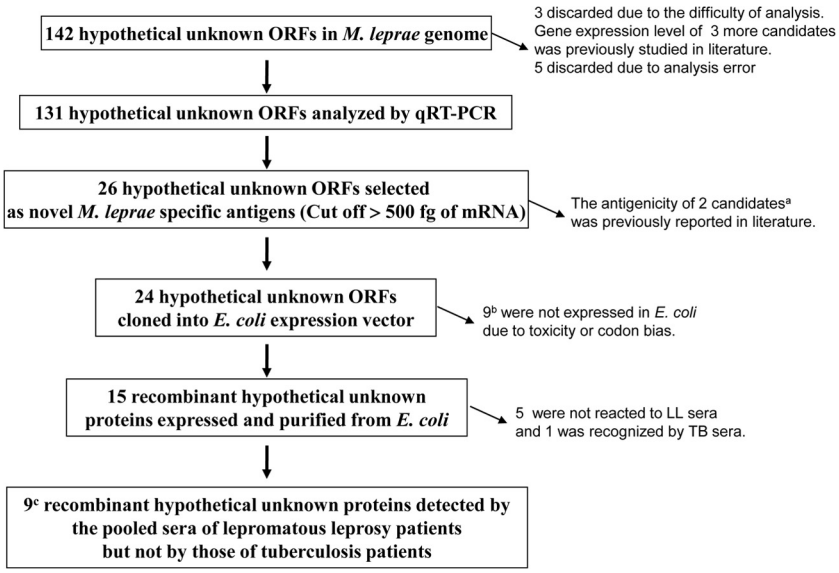

FIG 1 Flow chart of process to identify the 9 hypothetical unknowns in the present study. a, ML0573 and ML0574; b, ML0023, ML0070, ML0217, ML0614, ML0920, ML0959, ML1010, ML1575, and ML2630; c, ML0121, ML0188, ML0448, ML0527, ML0755, ML0953, ML2044, ML2313, and ML2666.

pooled from eight LL patients whose bacillus index (BI) was 6 and who were randomly selected. Five sera from cavitary TB patients were pooled to investigate their cross-reactivity to *M. leprae* recombinant antigens.

Western blot analysis. Quantities of proteins were measured by the bicinchoninic acid (BCA) assay (Pierce, Rockford, IL). Each novel recombinant antigen candidate (0.5 μg/lane) was subjected to electrophoresis in 15% SDS-PAGE gels and transferred onto nitrocellulose membranes. Blots were blocked with blocking buffer (3% bovine serum albumin [BSA] in phosphate-buffered saline [PBS]–0.05% Tween 80) for 2 h and then probed with both pools of diluted sera (1:5,000 dilution for leprosy sera and 1:1,000 dilution for TB sera). Blots were performed by probing with secondary anti-human IgG–alkaline phosphatase (Sigma, St. Louis, MO) and developed by using 5-bromo-4-chloro-3-indolyl phosphate–nitroblue tetrazolium (BCIP/NBT) (Sigma, St. Louis, MO) (23).

RESULTS

Establishment of the gene expression profile of *M. leprae*-specific hypothetical unknown ORFs and selection of novel antigen candidates. Previously, a global analysis of *M. leprae* transcripts by use of a microarray revealed that a few of the hypothetical unknown genes were expressed during infection (29). Although this global analysis has provided valuable information in bacterial pathogenesis studies (29), it is not sufficient to evaluate the transcriptional status of hypothetical unknown genes that are usually expressed at low levels.

In order to prove the existence of all class VI hypothetical unknowns in the *M. leprae* proteome, real-time PCR assays based on target-specific primers tagged with SYBR green I as fluorescent probes were performed to determine the relative quantification of hypothetical unknown ORFs in cDNAs synthesized from total *M. leprae* RNAs. SYBR green I binds to double-stranded DNA (dsDNA) generated with each progressive cycle of the PCR and emits a fluorescence signal which is quantitatively measured to track the amplification of DNA. There is a quantitative relationship between the amount of starting template and the PCR product at the exponential phase of the PCR (31). We established the gene expression profile of 131 hypothetical unknown ORFs which had not been studied previously. This target-based gene expression analysis revealed that the majority of the hypothetical unknown ORFs (60%) expressed less than 100 pg of mRNA (Fig. 1 and Table 1).

The transcript of ESAT-6/ML0049, which has proved to be an important T cell antigen (24), was expressed at low levels in both nude mouse-derived *M. leprae* and *M. leprae* recovered from MB patients (29). The gene expression levels of 26 hypothetical unknown genes were found to be in the range of 1 pg to 500 fg, similar to that of ESAT-6 (Table 2); the immunogenicity of ML0573 and ML0574 had already been evaluated in regions of endemicity (14). Thus, 24 of the 26 hypothetical unknowns were cloned, and attempts were made to express them in *E. coli* (Fig. 1).

It should be mentioned that the class VI hypothetical unknowns encoded by 142 theoretical ORFs had no orthologous genes in any of the mycobacterial genome databases (22, 26) at the time of initiation of this work. Subsequent advances in next-generation DNA sequencing technology resulted in the completion of a large number of microbial genome sequences, and advanced bioinformatic analysis has further refined the annotation of *M. leprae* ORFs over those from the previous genome sequences (32). In a reflection of this trend, the current list of genes encoding functional class VI proteins (http://www.pasteur.fr/recherche/unites/Lgmb/NATURE_DATA/ML_gene_list) has been adjusted continuously but has retained the majority of those in the previous list. However, in order to further evaluate whether these class VI hypothetical unknowns have homology to those expressed in other human pathogens, the 24 *M. leprae*-specific proteins were examined anew by BLASTP (http://www.ncbi.nlm.nih.gov/BLAST). A few unknowns appeared to contain orthologues among other pathogens but, again, mostly retained their status as hypothetical unknowns (data not shown). For instance, the original 26 hypothetical unknowns have been categorized into four subclasses based on E values (Table 3). In BLASTP, 9 of the 26 hypothetical unknowns showed some amino acid sequence similarity to hypothetical proteins in other mycobacterial genome da-

TABLE 1 qRT-PCR analysis of 131 hypothetical unknown ORFs

ORF	C_T value[a] for genomic DNA								C_T value[a] for cDNA	Level of gene expression
	1 ng	100 pg	10 pg	1 pg	500 fg	100 fg	50 fg	10 fg		
ML0009										
ML0840										
ML0939										
ML2045										
ML2669										
ML0464		22.3	25.8	29.5	32	33.5	34	36.4	36.7	<10 fg
ML0926		22.1		30	32	35.9	36.6	41	NA	<10 fg
ML1148	18.6	22.1	25.2	28.9	29.4	32.8	33.2	35.4	35.2	10 fg
ML0473		21.4	24.4	28.4	30.7	33.4	34	36.4	34.8	10–50 fg
ML0218		33.2	36.7	36.3	38.1	38.7	39	42.7	40.1	10–50 fg
ML0291		21.4	24.6	28.4	29.8	31.2	33	34.5	34.3	10–50 fg
ML0293[b]		21.2	24.8	28.8	30.4	31.1	32.7	36.8	34	10–50 fg
ML0938		21.9		28.9	30	31.3	32.2	34	33.4	10–50 fg
ML1001		23.4	27.3	32.6	34.5		36	38.7	36.5	10–50 fg
ML0325		25.6	27	32	33.1	34.7		37.9	35.4	10–100 fg
ML0927		26.4	30.8	33.9	34.2	37.5		41	38.3	10–100 fg
ML0152		22.1	26	29.5	30.4	32.5	34.1		35.6	<50 fg
ML0162		23.7	27.9	32.4	40.6	35	36.1		37	<50 fg
ML0470	17.4	21	24.6	28.7	29.7	31.7	32.1		36.4	<50 fg
ML0659		29	34	39.5		41.6	42.8		NA	<50 fg
ML1057							38.9		39.6	<50 fg
ML1210[b]	19.5	23.5	27.4	32	32.5	35.8	38.8		NA	<50 fg
ML1605	18.6	23.3	26.3	30.3	31.4	34			NA	<50 fg
ML1915	21.1	24.9	29	32.4	33.6	35.7	36.6		37	<50 fg
ML1976[b]	20.4	24.2	27.9	31.9	32.5	34.5	35.6		NA	<50 fg
ML1979	19.4	24	27.6	32.5	33	34.9	35.3		36.9	<50 fg
ML1989[b]	19.6	23.6	27.7	31.8	34	36			36.5	<50 fg
ML2013	19.7	23.8	27.9	31.1	33.2	35			NA	<50 fg
ML2346				33.5	34.1	36.1	36.5	NA	38	<50 fg
ML2476[b]	18.8	23.8	26.7	30.2	31.2	33.6	33.7		34.2	<50 fg
ML2478	18.3	23.8	26.9	31.4	33.1	35.7	35		NA	<50 fg
ML2491	24	29.5	33.5	37.9	39.5	NA	41.5		41.9	<50 fg
ML2562	20.5	24.7	29.3	32.1	33.2	35	36.1		37.1	<50 fg
ML2603	23	30	35.6	39	40.5	41.9	42.6		NA	<50 fg
ML2629	18.7	24.6	28.7	33.5	33.1	37	36.9		NA	<50 fg
ML0024		24.6	28.3	31.6	NA	37.6	39	NA	38.1	50 fg
ML0664		22.4	26.1	30.4		34.1	34.3		34.5	50 fg
ML0777		22.1	26	29.5	30.8	32.8	35		35.5	50 fg
ML1188	18.5	22.5	26.1	30	30.5	32.9	34.3		34.5	50 fg
ML1189		23.7	27.9	32.1	33	36.1	37		36.8	50 fg
ML1292	17.7	21.4	25.1	29	30.2	32	33		33.6	50 fg
ML1420[b]	19	22.9	26.5	30.1	31.9	34.1	34.7		39.8	50 fg
ML1761	18.1		25.8	29.7	30.1	31.9	32.9		32.6	50 fg
ML2091	19	23.4	27	31.2	31.8	32	34.2		34.2	50 fg
ML2170	23.5	28.5	31.9	36.1	36.8	38.5	39.3	41.2	39.4	50 fg
ML2172	22	26.4	31.7	33.7	36	38.5	39	43	39.6	50 fg
ML2497	20.9	26.7	30.3	34.8	36	39	40.5		40.4	50 fg
ML0025			35	37.1	38	39.4	42.5		41.1.	50–100 fg
ML0126				32.7	33	34.6	36.8	NA	35.6	50–100 fg
ML0292[b]		20.8	24.4	27.7	28.7	31.9	32.6	34.6	31	50–100 fg
ML0369		23.3	26.1	30.6	32.1	33.4	36.4		34.5	50–100 fg
ML0863		22.7	26	29	31	32	34		33.2	50–100 fg
ML0957		23	25.9	29.7	32	32.8	35.3		34.7	50–100 fg
ML1018	18.7	22.4	25.7	27.9	30.2	31.6	34		32.5	50–100 fg
ML1243[b]	23.9	29	33.2	38.5	39.1	40.8	42.5		42.3	50–100 fg
ML1275	18.8	22.9	26.7	31.3	31.7	33.4	36		34.1	50–100 fg
ML1523	18.6	22.4	26.7	29.8	30.8	33.4	34.7		34.5	50–100 fg
ML1602	18	21.9	25.4	29.2	30.2	31.5	34.3	36.5	33.2	50–100 fg
ML1763	18.9		26.9	30.5	31	33.5	36		34	50–100 fg
ML1788[b]	18.8	22.5	26	30.2	31.2	33.2	34.8	39	34	50–100 fg

(Continued on following page)

TABLE 1 (Continued)

ORF	C_T value[a] for genomic DNA								C_T value[a] for cDNA	Level of gene expression
	1 ng	100 pg	10 pg	1 pg	500 fg	100 fg	50 fg	10 fg		
ML1829	17.8	22.5	26.3	29.8	30.5	31	32.7		32.5	50–100 fg
ML2201	18.9	23.3	27.1	31	31.2	32.3	32.8		32.6	50–100 fg
ML0950		23.8	28	31.1	32	33.6			35.7	<100 fg
ML0964		25.3	29.5	32.8	33.4	35			36.6	<100 fg
ML0950		23.8	28	31.1	32	33.6			35.7	<100 fg
ML1119	20.3	24.3	29.2	33.2	33.2	36			NA	<100 fg
ML1517	21.4	25.9	30.9	34.2		36.4			38.8	<100 fg
ML1604	17.9	22	25.5	29.6	30.2	32.8			34.2	<100 fg
ML1717[b]	18.8		26.2	30.6	31.1	33.3			35	<100 fg
ML1821	18	22.4	25.7	29.7	33.6	34.4			40.7	<100 fg
ML2452	19	24	27.9	32	32.6	34.6			35.2	<100 fg
ML2265	17.6	21.6	25.1	29	29.4	NA	NA		31.3	<100 fg
ML2178	21.4	25.9	31.2	34.6	35.9	37			NA	<100 fg
ML0394				34.8	NA	38.1	NA	39.4	38.5	100 fg
ML0663		23.2	26.2	31	31.9	33			32.9	100 fg
ML0949		22.6	31	32.1	35	36.1			36.8	100 fg
ML1011	18	21.8	25	28.7	30.2	32.4	32.8	35	32	100 fg
ML1106	19	22.8	26.4	29.9	31.3	33.6		38.8	33.3	100 fg
ML1344[b]	17.5	21.4	24.7	28.5	29.3	31.8	NA		34.3	100 fg
ML2158	18.9	24.1	27.6	31.5	32.7	33.2	35		33.6	100 fg
ML2264	18.8	23.6	27.3	32	33.3	35.4	36.2		35.5	100 fg
ML2468	18	22.1	25.2	29.5	30	32.5	34.4		32.1	100 fg
ML0472[b]		21.3	25.3	28.2	31.9	32.8		34.4	32	100–500 fg
ML0656		22.4	25.8	30	30.8	34.3	35		33.6	100–500 fg
ML0679	18.5	22.4	25.9	29.5	30.5	33.4	33.8		31	100–500 fg
ML0947		21.9	29.1	30.7	31.8	32.5	33.5	39	32	100–500 fg
ML1186	18.5	22.3	25.7	29.3	31	35.1			34.5	100–500 fg
ML1294	17.2	21.4	24.6	28.1	29.3	31.7	33.2		31	100–500 fg
ML1572[b]		21.5	25.8	29.7	30.4	32.6	33.5		31.7	100–500 fg
ML1601	18	22.2	25.3	29.2	31	34.3			32.1	100–500 fg
ML1603	18.1	22	25.7	28.7	29.4	32.1			31	100–500 fg
ML1793	18.7	21.9	26.5	30.8	31	32.6	34.5		31.8	100–500 fg
ML1796	18.3	22.8	26	30.4	32.2	32.8	33.6		32.5	100–500 fg
ML1928	17.5	21.8	25	28.4	29.1	31.8	32.7		30.9	100–500 fg
ML1972	19.3	23.4	26.6	30.6	32.1	34.4	35.5		33.3	100–500 fg
ML2035[b]	22.5	26.3	29.6	33.3	33.8	36			35	100–500 fg
ML2176	19.2	23.9	28.4	31.5	32.5	34.2	38.7	39.8	32.5	100–500 fg
ML2252	19.9	24.3	27.2	32	32.8	36	37.9	41.1	35.5	100–500 fg
ML2288	19.4	24	27	31.8	32.7	36.2			34.5	100–500 fg
ML2379	18.8	24.3	27.7	32	32.4	36.7	37		33.7	100–500 fg
ML2621	19.6	25.3	29.5	33.7	36	37.9	39		36.8	100–500 fg
ML0963		23.4	29.8	32.7	35				n/a	<500 fg
ML0023[b]		24.6	26.9	38.9	37.1	37.9	38.6	NA	37.4	500 fg
ML0121		23.4	26.9	31.8	32.9	33.5	36.2		31.8	500 fg
ML0127		25.3	29.9	32.7	33.2	34.3	NA	NA	33.5	500 fg
ML0265[b]		29.3	32.1		35	39.1	42.7		35.4	500 fg
ML0757		23	26.7	30	31.8	33.1	35.4		31.7	500 fg
ML0928		28.8	34	39	40	42			40.3	500 fg
ML1445	19	22.7	26.3	30	32	34	36		32.4	500 fg
ML2283	17.7	22.7	25.6	29.7	31.1	34.9			31	500 fg
ML2284	19.2	24	31.6	32	34.3	35.2			34.2	500 fg
ML0448		21.7	25.2	29.1	33.3	34.2			31.4	500 fg-1 pg
ML0573		21.6	24.7	28.8	33.4	34.8	34	41	32	500 fg-1 pg
ML0755		23.2	26.7	31	31.8	33.1	35.4		31.7	500 fg-1 pg
ML0920		23.8	27.4	31.1	32.5	35			31.5	500 fg-1 pg
ML0953		27	29.5	33.8	35.1			39.2	34.1	500 fg-1 pg
ML1949	19.6	23.7	27.1	31.4	32.9	34.1			32	500 fg-1 pg
ML2630	17.3	22.8	26.5	30.9	31.6	32.9	NA		31.3	500 fg-1 pg
ML0141[b]			36	37.6	39.2	40.6	42.7	NA	36.5	1 pg
ML0188		23.9	29.1		35.4	36.4	36.8	NA	34.6	1 pg

(Continued on following page)

TABLE 1 (Continued)

ORF	C_T value[a] for genomic DNA								C_T value[a] for cDNA	Level of gene expression
	1 ng	100 pg	10 pg	1 pg	500 fg	100 fg	50 fg	10 fg		
ML0217[b]		23.8	28.4	33.9	36.9		37.1	40.4	32	1 pg
ML0070		25.2	30.7	33.6	33.5	34	34.2	35.4	33.6	1 pg
ML0574		23.1	27	30.5	31	32	33.2		30.9	1 pg
ML0588		22.8	26	30.5	31.8	32.7	34.8		30.1	1 pg
ML0614	20.8	25	29.1	32.4	36	37.5			32	1 pg
ML0959		24	29	32.5	33.4	36.2			32.6	1 pg
ML1010	18.4	22.4	25.2	28.7	30.2	32.2	32.9		28.5	1 pg
ML1384	18.8	22.7	25.2	30.4	31	34.8	35.2		30.7	1 pg
ML1575	26.4	31.1	34.4	37.8	38.9	41.4			37.9	1 pg
ML2044	21	25.7	30	33.7	34.7	37	39.1		33.9	1 pg
ML2307	19.9	24.6	27.2	31.8	33.2	34.8	35.2		31.5	1 pg
ML2313	20.9	26	30.5	34.3	35.8	38.7			32.3	1 pg
ML2651	19.3	23.1	27.3	30.9	32	35.7	NA		31.1	1 pg
ML2666	19.5	25.1	28	32	32.3	34.7	35.6		32.1	1 pg
ML0527		22.4	25.4	29.2	30.2	32.4	33.2		28.9	>1 pg
ESAT-6		22	26	29.4	31.7	32.4	35.2		29	≈1 pg
ML0380		28	31.5	36.7					32.1	1–10 pg
ML2038		20.2	24	27.5	31	32.9	34		26.2	1–10 pg

[a] The C_T (cycle threshold) value is the number of cycles at which the fluorescence signal generated by PCR just exceeds the background fluorescence level (threshold) reached at the exponential phase of the PCR (http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_042502.pdf). NA, not applicable.
[b] Recent gene annotation analysis revealed that the gene is a pseudogene or doubtful CDS. The expression level of each gene was determined by comparison to the results of qRT-PCR assays performed using a series of genomic DNA standards.

tabases (Table 3). The lengths of amino acid sequences comparable to those of the 26 unknowns were mostly less than 100 amino acids, with high gap scores and identities of <25%, indicating that the selected antigen candidates do not have orthologues in other pathogens and are considered to be *M. leprae* specific (33).

Purification and serological reactivity of recombinant *M. leprae*-specific antigens. Fifteen of the 24 new candidates were

TABLE 2 qRT-PCR analysis of 26 selected hypothetical unknown ORFs

ORF	C_T value for genomic DNA[a]							C_T value for cDNA	Level of gene expression[b]
	100 pg	10 pg	1 pg	500 fg	100 fg	50 fg	10 fg		
ML0023	24.6	26.9	38.9	37.1	37.9	38.6	NA	37.4	500 fg
ML0070	25.2	30.7	33.6	33.5	34	34.2	35.4	33.6	1 pg
ML0121	23.4	26.9	31.8	32.9	33.5	36.2		31.8	500 fg
ML0141		36	37.6	39.2	40.6	42.7	NA	36.5	1 pg
ML0188	23.9	29.1		35.4	36.4	36.8	NA	34.6	1 pg
ML0217	23.8	28.4	33.9	36.9		37.1	40.4	32	1 pg
ML0448	21.7	25.2	29.1	33.3	34.2			31.4	500 fg-1 pg
ML0527	22.4	25.4	29.2	30.2	32.4	33.2		28.9	>1 pg
ML0573	21.6	24.7	28.8	33.4	34.8	34	41	32	500 fg-1 pg
ML0574	23.1	27	30.5	31	32	33.2		30.9	1 pg
ML0588	22.8	26	30.5	31.8	32.7	34.8		30.1	1 pg
ML0614	25	29.1	32.4	36	37.5			32	1 pg
ML0755	23.2	26.7	31	31.8	33.1	35.4		31.7	500 fg-1 pg
ML0920	23.8	27.4	31.1	32.5	35			31.5	500 fg-1 pg
ML0953	27	29.5	33.8	35.1			39.2	34.1	500 fg-1 pg
ML0959	24	29	32.5	33.4	36.2			32.6	1 pg
ML1010	22.4	25.2	28.7	30.2	32.2	32.9		28.5	1 pg
ML1384	22.7	25.2	30.4	31	34.8	35.2		30.7	1 pg
ML1575	31.1	34.4	37.8	38.9	41.4			37.9	1 pg
ML1949	23.7	27.1	31.4	32.9	34.1			32	500 fg-1 pg
ML2044	25.7	30	33.7	34.7	37	39.1		33.9	1 pg
ML2307	24.6	27.2	31.8	33.2	34.8	35.2		31.5	1 pg
ML2313	26	30.5	34.3	35.8	38.7			32.3	1 pg
ML2630	22.8	26.5	30.9	31.6	32.9	NA		31.3	500 fg-1 pg
ML2651	23.1	27.3	30.9	32	35.7	NA		31.1	1 pg
ML2666	25.1	28	32	32.3	34.7	35.6		32.1	1 pg

[a] NA, not applicable.
[b] The transcription levels of 26 hypothetical unknown ORFs were similar to that of ML0049, an immunologically important antigen expressed at fairly low levels (22, 29).

TABLE 3 Amino acid sequence similarity of selected *M. leprae*-specific proteins[a]

Protein subclass	ORF	Molecular mass (kDa)	Sequence similarity (E value)	Mycobacterial protein(s) with amino acid sequence similarity[b]
VI.a	ML0070	9.1	None	NA
	ML0141	9.3	None	NA
	ML0448	10	None	NA
	ML0527	8.7	None	NA
	ML0574	11.4	None	NA
	ML0959	13.6	None	NA
	ML1010	8.4	None	NA
	ML1384	12.2	None	NA
	ML1575	11.1	None	NA
	ML2044	7.9	None	NA
	ML2651	11.6	None	NA
	ML2666	8.8	None	NA
VI.b	ML0121	9.6	3e−08	Hypothetical proteins of *M. paratuberculosis*
	ML0023	11.7	3e−06	
	ML0188	9.2	3e−08	Conserved hypothetical unknown of *M. paratuberculosis*
	ML0217	8.4	1e−06	
	ML0573	9.5	1e−05	Hypothetical protein of *M. kansasii*
	ML0588	8.5	3e−04	PPE protein of *M. tuberculosis*
	ML0953	8.6	1e−06	Putative nucleic acid binding protein of *M. kansasii*
VI.c	ML0614	10.2	3e−17	Hypothetical proteins of *M. tuberculosis* T46
	ML2630	13	7e−10	Conserved membrane protein of *M. tuberculosis* T92
	ML1949	12.5	5e−17	Hypothetical proteins of *M. kansasii*
VI.d	ML0755	9.6	1e−22	Hydrolase of *M. kansasii*
	ML0920	6.3	1e−102	Hypothetical protein of *M. tuberculosis* T17
	ML2307	9.3	4e−52	Transcription factor WhiB4 of *M. paratuberculosis*
	ML2313	21.8	3e−71	Transcriptional regulator (PadR) of *M. avium* 104

[a] All ORFs in subclass VI.a appear to have no homology. *M. leprae* proteins with BLASTP values between 1×10^{-4} and 1×10^{-8} are considered to have low homology (subclass VI.b). Proteins with E values between 1×10^{-8} and 1×10^{-20} have moderate homology (subclass VI.c), and those with E values of $<1 \times 10^{-20}$ have high homology (subclass VI.c).

[b] Most similar one among several proteins found in BLASTP. NA, not applicable.

successfully expressed as His-tagged fusion proteins in the *E. coli* T7 promoter-driven vector system and were purified by immobilized-metal affinity chromatography (Fig. 2A). In order to evaluate the true immunogenicity of these novel antigens, we performed Western blot analysis using sera from both LL and cavitary TB patients (Fig. 2B and C).

The seroreactivities to all recombinant hypothetical unknown proteins showed that 10 of the 15 recombinant antigens were recognized by circulating antibodies in the pooled sera from LL patients (Fig. 2B). Western blot analysis using pooled sera from five cavitary TB patients detected a single recombinant antigen among the 15 proteins, namely, ML2307, which had high homology with orthologues of other bacteria (Table 3 and Fig. 2C). Although all ORFs of subclass VI.d showed higher homology with the orthologues of *M. tuberculosis* than with other candidate antigens (Table 3), sera from cavitary TB patients did not react to ML0755, ML0920, and ML2313 in this analysis (Fig. 2C). ML0141, ML0588, ML1384, and ML2651, in subclass VI.a, and ML1949, in subclass VI.c, were not seroreactive against LL or TB patient sera (Fig. 2B and C). Therefore, ML0121, ML0188, ML0448, ML0527, ML0755, ML0953, ML2044, ML2313, and ML2666 induced *M. leprae*-specific immune responses.

DISCUSSION

The overall objective of this work was to address gene expression levels of all *M. leprae*-specific hypothetical unknowns and to evaluate antigen immunogenicity. Eventually, the selected antigens in this study will allow detection of individuals exposed to and/or infected by *M. leprae* in regions where leprosy is endemic.

Previously, 31 hypothetical unknowns, including some of those addressed in the present study, as well as their corresponding peptides, were expressed and their antigenic potentials for early diagnosis of leprosy evaluated using the IGRA format in various regions of endemicity (6, 7, 14, 16, 22). Some of these *M. leprae*-specific antigens were shown to be valuable diagnostic reagents for specific diagnosis of those infected with *M. leprae*, particularly occupational contacts or HHCs of MB or PB patients (14, 15, 17, 22). However, most of the hypothetical unknown proteins and peptides were recognized by a considerable number of endemic controls and TB patients and showed apparent variation in reactivity to racially and geographically different study populations (14–17, 22). Although some of the antigens induced low levels of IFN-γ production in some individuals in the endemic control group who had strongly responded to *M. leprae* lysate, the level (100 to 200 pg) of released IFN-γ in these individuals was just above the cutoff value (100 pg), making it difficult to determine T cell positivity in individuals in the endemic control group with *M. leprae* exposure (16, 17). The question therefore arose as to which, if any, of these proteins are expressed in *M. leprae*.

The genome of *M. leprae* contains an exceptionally large number of pseudogenes and genes for hypothetical unknowns (26). These hypothetical unknowns (class VI) are mostly low-molecular-weight proteins (14, 22). Therefore, it is possible that their ORFs are not genuine or are barely expressed and available for T or B cell recognition. Since the identification of these antigens was highly dependent on *in silico* analysis, low availability to the human immune system and HLA variability might be responsible for the variable reactivity and low sensitivity of IGRAs for different ethnic populations (14–16, 22). Therefore, a major remaining challenge in the development of leprosy diagnostic assays was to select the hypothetical unknown antigens that are definitely expressed and evoke a more definite immune response in individuals with *M. leprae* infection/exposure, regardless of genetic or geographic situation, and can maintain the specificity of some current IGRAs. In this study, we attempted to establish the gene transcriptional levels of 131 *M. leprae*-specific proteins (class VI) to enhance the screening process for novel antigens, using target-based

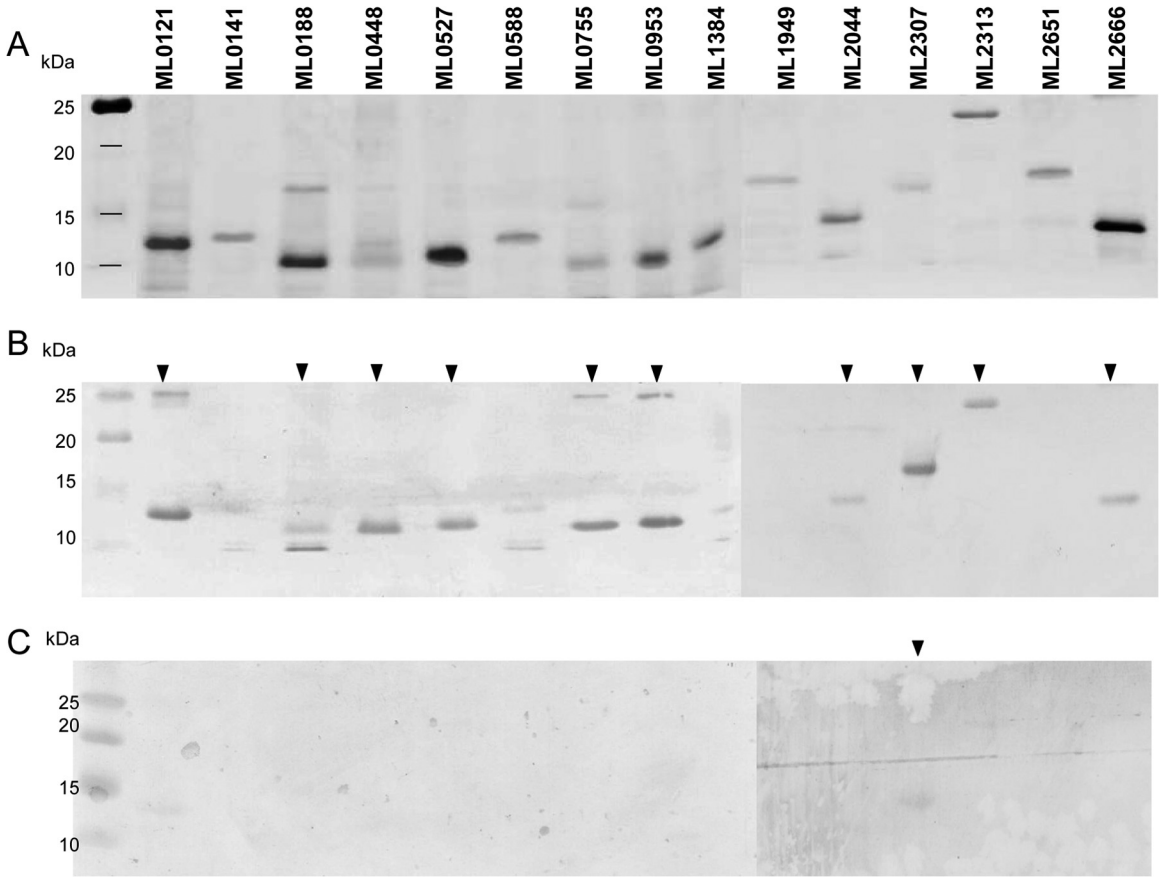

FIG 2 Western blot analysis of recombinant forms of selected hypothetical unknown proteins. (A) SDS-PAGE gel with silver nitrate staining after SDS-PAGE analysis of the recombinant candidates. (B) Western blot hybridized with pooled sera from LL patients (1:5,000 dilution) whose BI was around 6. (C) Western blot from identical gel hybridized with pooled sera from cavitary TB patients. Filled arrowheads indicate candidates that reacted with LL patient sera and TB patient sera.

qRT-PCR assays (Table 1). Of the class VI genes, 26 of the hypothetical unknown ORFs were selected as actively expressed genes with high positive PCR signals (>500 pg of cDNA) (Table 2). Only two of the selected antigens, ML0573 and ML0574, had already been studied as T cell antigens and reported in the literature, but these failed to induce *M. leprae*-specific T cell immunity in a Brazilian population (14, 15).

Recently, an IGRA validated that five of the most promising recombinant hypothetical unknowns (ML0126, ML1420, ML1989, ML2283, and ML2346) and 22 of their peptides induced IFN-γ secretion from PB patients or HHCs in areas where leprosy is endemic in Brazil, Bangladesh, Nepal, Pakistan, and Ethiopia (16). In agreement with previous studies (14, 15, 17, 22), the ML2283 recombinant protein and ML2283-derived peptides induced the highest *M. leprae*-specific T cell responses (16). However, none of the Brazilian subjects exhibited T cell responses to any of the peptide antigens, and the ML1420-derived peptides induced *M. leprae*-specific T cell responses only in the Pakistani population (16).

In the present study, the gene expression level of ML2283 was found to be 500 fg, which was relatively high compared to those of the other hypothetical unknowns. The ORFs of ML0126, ML1420, ML1989, and ML2346, which induced various responses among three populations, were expressed at 100 fg (Table 1). The literature and our study show that the low sensitivity of IGRA to current peptide antigens may be attributable to low antigen availability and poor T cell recognition of the parent proteins by the host immune system.

A number of well-known *M. leprae* antigens (MMPI, MMPII, GroES, ML0049, and ML0050) induce strong cell-mediated and humoral immune responses in leprosy patients (24, 34–38). Araoz et al. (6, 7) showed that a few of the recombinant hypothetical unknown proteins can be recognized by circulating antibodies in sera of leprosy patients but also induce *M. leprae*-specific T cell responses in leprosy and endemic control populations. Surprisingly, our study showed that the majority of the selected antigen candidates were detected by circulating antibodies in sera of LL patients, suggesting that these antigens can be processed and accommodated on the major histocompatibility complex (MHC) molecule to present to T helper cells (Fig. 2B). Although all cavitary TB patients presented high bacterial loads, seroreactivities of these patients were very low toward most of the antigen candidates (Fig. 2C). Therefore, these novel recombinant antigens have considerable potential for improving the sensitivity and specificity of leprosy diagnosis.

The dramatic advance of bioinformatics revealed that 17 ORFs among a total of 142 hypothetical unknown ORFs contain doubtful coding DNA sequences or are pseudogenes

(http://genolist.pasteur.fr/Leproma/) (Table 1). Of nine *M. leprae*-specific proteins in the present study, ML2044 was found to be encoded by a pseudogene coding for a small polypeptide. Recently, a proteomic study by de Souza et al. (39) showed that five proteins encoded by pseudogenes are expressed in the *M. leprae* proteome, but these were not present among the 26 selected proteins. Also, transcriptional analysis by others (32) showed that a considerable number of pseudogenes are expressed. Therefore, it is important to investigate whether ML2044, as a pseudogene product, plays a role in host-pathogen interaction and may be a potential antigen in leprosy diagnosis.

Bacterial proteins are distinctive in their subcellular locations, and secreted or envelope-associated proteins are generally agreed to be crucial determinants of host immunopathogenesis and prospective antigens in disease diagnosis and vaccine studies. In this light, we performed bioinformatic analysis of all hypothetical unknowns by using the SignalP 4.0 (http://www.cbs.dtu.dk/services/SignalP/) and TMHMM 2.0 (40) programs, seeking evidence of secretion signals or transmembrane domains. No evidence of secretory signals was found in the hypothetical proteins, but transmembrane domains appeared in ML0188, suggesting that this protein may be recognized efficiently by the host immune system. Interestingly, ML0588, which was recognized by LL sera, showed amino acid sequence similarity to *M. tuberculosis* Rv1387, containing a PPE repeated motif (http://genolist.pasteur.fr/Leproma/). In *M. tuberculosis*, proteins in the PPE family contain conserved proline and glutamate (PE and PPE) motifs and mostly appear in the cell envelope (41). The PE and PPE proteins are known to be involved in antigenic variation and stimulate a robust cellular immunity via Toll-like receptors (TLRs), contributing to mycobacterial pathogenesis (41, 42). Although the overlap sequence between ML0588 and Rv1387 was confirmed to contain only 38 amino acids, it is still possible that ML0588 induces *M. leprae*-specific responses and plays a role in *M. leprae* virulence and leprosy diagnosis.

In the development of diagnostic tests using *M. leprae*-specific antigens, the ultimate goal is to detect *M. leprae*-infected individuals with asymptomatic leprosy as the biological reservoir and to achieve the goal of disease eradication. Until now, a CMI-based test using some of *M. leprae*'s unique proteins provided the greatest potential to detect HHCs, who are exposed to and/or infected by *M. leprae* (14–18, 22) but do not exhibit obvious clinical symptoms, as well as TT/PB patients. However, several studies showed that most of the selected proteins or peptide antigens could not achieve a reasonable sensitivity of IGRAs, particularly in regions where leprosy is endemic (14–16, 22). Thus, such assays are limited in their use for diagnosis of human leprosy in routine clinical practice. In the present study, we demonstrated that 9 selected antigen candidates are likely expressed in *M. leprae* and may reasonably be expected to enhance the sensitivity of immunological assays for early diagnosis of leprosy.

A major concern of CMI-based diagnostic assays using peptide antigens is the variable immune response in genetically different populations. In order to circumvent this problem, we believe that a mixture of peptides binding all major HLA-DR types will provide sufficient coverage of all haplotypes of different populations with *M. leprae* infection and/or exposure. Recently, a consensus prediction method for 11 different major HLA-DR alleles was developed to combine three top-performing MHC class II peptide binding prediction methods and was shown to achieve the best overall performance (43). In future, the identification of *M. leprae*-specific peptides (from these novel candidates) that are restricted to all major HLA-DR types and specifically activate the maximal IFN-γ response in peripheral blood mononuclear cells (PBMCs) from PB patients and HHCs should allow improvements in the sensitivity and specificity of IGRAs for early diagnosis of leprosy.

ACKNOWLEDGMENTS

This study was supported by NIH, NIAID, grant RO1 AI-082575-01, NIH, NIAID, grant R01 AI-047197, and NIH, NIAID, contract N01 AI-025469.

REFERENCES

1. **Britton WJ, Lockwood DN.** 2004. Leprosy. Lancet **363**:1209–1219.
2. **Duthie MS, Goto W, Ireton GC, Reece ST, Cardoso LP, Martelli CM, Stefani MM, Nakatani M, de Jesus RC, Netto EM, Balagon MV, Tan E, Gelber RH, Maeda Y, Makino M, Hoft D, Reed SG.** 2007. Use of protein antigens for early serological diagnosis of leprosy. Clin. Vaccine Immunol. **14**:1400–1408.
3. **Sekar B.** 2007. Recent advances in immunodiagnosis of leprosy. Indian J. Lepr. **79**:85–106.
4. **Weir RE, Brennan PJ, Butlin CR, Dockrell HM.** 1999. Use of a whole blood assay to evaluate *in vitro* T cell responses to new leprosy skin test antigens in leprosy patients and healthy subjects. Clin. Exp. Immunol. **116**:263–269.
5. **Weir RE, Fine PE, Floyd S, Stenson S, Stanley C, Branson K, Britton WJ, Huygen K, Singh M, Black G, Dockrell HM.** 2008. Comparison of IFN-gamma responses to mycobacterial antigens as markers of response to BCG vaccination. Tuberculosis (Edinb.) **88**:31–38.
6. **Araoz R, Honore N, Banu S, Demangel C, Cissoko Y, Arama C, Uddin MK, Hadi SK, Monot M, Cho SN, Ji B, Brennan PJ, Sow S, Cole ST.** 2006. Towards an immunodiagnostic test for leprosy. Microbes Infect. **8**:2270–2276.
7. **Araoz R, Honore N, Cho S, Kim JP, Cho SN, Monot M, Demangel C, Brennan PJ, Cole ST.** 2006. Antigen discovery: a postgenomic approach to leprosy diagnosis. Infect. Immun. **74**:175–182.
8. **Brennan PJ.** 2000. Skin test development in leprosy: progress with first-generation skin test antigens, and an approach to the second generation. Lepr. Rev. **71**(Suppl):S50–S54.
9. **Buhrer-Sekula S, Smits HL, Gussenhoven GC, van Leeuwen J, Amador S, Fujiwara T, Klatser PR, Oskam L.** 2003. Simple and fast lateral flow test for classification of leprosy patients and identification of contacts with high risk of developing leprosy. J. Clin. Microbiol. **41**:1991–1995.
10. **Cho SN, Yanagihara DL, Hunter SW, Gelber RH, Brennan PJ.** 1983. Serological specificity of phenolic glycolipid I from *Mycobacterium leprae* and use in serodiagnosis of leprosy. Infect. Immun. **41**:1077–1083.
11. **Dockrell HM, Brahmbhatt S, Robertson BD, Britton S, Fruth U, Gebre N, Hunegnaw M, Hussain R, Manandhar R, Murillo L, Pessolani MC, Roche P, Salgado JL, Sampaio E, Shahid F, Thole JE, Young DB.** 2000. A postgenomic approach to identification of *Mycobacterium leprae*-specific peptides as T-cell reagents. Infect. Immun. **68**:5846–5855.
12. **Duthie MS, Hay MN, Morales CZ, Carter L, Mohamath R, Ito L, Oyafuso LK, Manini MI, Balagon MV, Tan EV, Saunderson PR, Reed SG, Carter D.** 2010. Rational design and evaluation of a multiepitope chimeric fusion protein with the potential for leprosy diagnosis. Clin. Vaccine Immunol. **17**:298–303.
13. **Duthie MS, Ireton GC, Kanaujia GV, Goto W, Liang H, Bhatia A, Busceti JM, MacDonald M, Neupane KD, Ranjit C, Sapkota BR, Balagon M, Esfandiari J, Carter D, Reed SG.** 2008. Selection of antigens and development of prototype tests for point-of-care leprosy diagnosis. Clin. Vaccine Immunol. **15**:1590–1597.
14. **Geluk A, Klein MR, Franken KL, van Meijgaarden KE, Wieles B, Pereira KC, Buhrer-Sekula S, Klatser PR, Brennan PJ, Spencer JS, Williams DL, Pessolani MC, Sampaio EP, Ottenhoff TH.** 2005. Postgenomic approach to identify novel *Mycobacterium leprae* antigens with potential to improve immunodiagnosis of infection. Infect. Immun. **73**:5636–5644.
15. **Geluk A, Ottenhoff TH.** 2006. HLA and leprosy in the pre and postgenomic eras. Hum. Immunol. **67**:439–445.
16. **Geluk A, Spencer JS, Bobosha K, Pessolani MC, Pereira GM, Banu S, Honore N, Reece ST, MacDonald M, Sapkota BR, Ranjit C, Franken**

KL, Zewdie M, Aseffa A, Hussain R, Stefani MM, Cho SN, Oskam L, Brennan PJ, Dockrell HM. 2009. From genome-based in silico predictions to ex vivo verification of leprosy diagnosis. Clin. Vaccine Immunol. **16**:352–359.

17. **Geluk A, van der Ploeg J, Teles RO, Franken KL, Prins C, Drijfhout JW, Sarno EN, Sampaio EP, Ottenhoff TH.** 2008. Rational combination of peptides derived from different *Mycobacterium leprae* proteins improves sensitivity for immunodiagnosis of *M. leprae* infection. Clin. Vaccine Immunol. **15**:522–533.
18. **Geluk A, van der Ploeg-van Schip JJ, van Meijgaarden KE, Commandeur S, Drijfhout JW, Benckhuijsen WE, Franken KL, Naafs B, Ottenhoff TH.** 2010. Enhancing sensitivity of detection of immune responses to *Mycobacterium leprae* peptides in whole-blood assays. Clin. Vaccine Immunol. **17**:993–1004.
19. **Geluk A, van Meijgaarden KE, Franken KL, Subronto YW, Wieles B, Arend SM, Sampaio EP, de Boer T, Faber WR, Naafs B, Ottenhoff TH.** 2002. Identification and characterization of the ESAT-6 homologue of *Mycobacterium leprae* and T-cell cross-reactivity with *Mycobacterium tuberculosis.* Infect. Immun. **70**:2544–2548.
20. **Geluk A, van Meijgaarden KE, Franken KL, Wieles B, Arend SM, Faber WR, Naafs B, Ottenhoff TH.** 2004. Immunological crossreactivity of the *Mycobacterium leprae* CFP-10 with its homologue in *Mycobacterium tuberculosis.* Scand. J. Immunol. **59**:66–70.
21. **Shepard CC, Draper P, Rees RJ, Lowe C.** 1980. Effect of purification steps on the immunogenicity of *Mycobacterium leprae.* Br. J. Exp. Pathol. **61**:376–379.
22. **Spencer JS, Dockrell HM, Kim HJ, Marques MA, Williams DL, Martins MV, Martins ML, Lima MC, Sarno EN, Pereira GM, Matos H, Fonseca LS, Sampaio EP, Ottenhoff TH, Geluk A, Cho SN, Stoker NG, Cole ST, Brennan PJ, Pessolani MC.** 2005. Identification of specific proteins and peptides in *Mycobacterium leprae* suitable for the selective diagnosis of leprosy. J. Immunol. **175**:7930–7938.
23. **Spencer JS, Kim HJ, Wheat WH, Chatterjee D, Balagon MV, Cellona RV, Tan EV, Gelber R, Saunderson P, Duthie MS, Reece ST, Burman W, Belknap R, Mac Kenzie WR, Geluk A, Oskam L, Dockrell HM, Brennan PJ.** 2011. Analysis of antibody responses to *Mycobacterium leprae* phenolic glycolipid I, lipoarabinomannan, and recombinant proteins to define disease subtype-specific antigenic profiles in leprosy. Clin. Vaccine Immunol. **18**:260–267.
24. **Spencer JS, Marques MA, Lima MC, Junqueira-Kipnis AP, Gregory BC, Truman RW, Brennan PJ.** 2002. Antigenic specificity of the *Mycobacterium leprae* homologue of ESAT-6. Infect. Immun. **70**:1010–1013.
25. **Manandhar R, LeMaster JW, Butlin CR, Brennan PJ, Roche PW.** 2000. Interferon-gamma responses to candidate leprosy skin-test reagents detect exposure to leprosy in an endemic population. Int. J. Lepr. Other Mycobact. Dis. **68**:40–48.
26. **Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG.** 2001. Massive gene decay in the leprosy bacillus. Nature **409**:1007–1011.
27. **Vissa VD, Brennan PJ.** 2001. The genome of *Mycobacterium leprae*: a minimal mycobacterial gene set. Genome Biol. **2**:REVIEWS1023.
28. **Williams DL, Oby-Robinson S, Pittman TL, Scollard DM.** 2003. Purification of *Mycobacterium leprae* RNA for gene expression analysis from leprosy biopsy specimens. Biotechniques **35**:534–541.
29. **Williams DL, Torrero M, Wheeler PR, Truman RW, Yoder M, Morrison N, Bishai WR, Gillis TP.** 2004. Biological implications of *Mycobacterium leprae* gene expression during infection. J. Mol. Microbiol. Biotechnol. **8**:58–72.
30. **Ridley DS, Jopling WH.** 1966. Classification of leprosy according to immunity. A five-group system. Int. J. Lepr. Other Mycobact. Dis. **34**:255–273.
31. **Malinen E, Kassinen A, Rinttila T, Palva A.** 2003. Comparison of real-time PCR with SYBR green I or 5′-nuclease assays and dot-blot hybridization with rDNA-targeted oligonucleotide probes in quantification of selected faecal bacteria. Microbiology **149**:269–277.
32. **Williams DL, Slayden RA, Amin A, Martinez AN, Pittman TL, Mira A, Mitra A, Nagaraja V, Morrison NE, Moraes M, Gillis TP.** 2009. Implications of high level pseudogene transcription in *Mycobacterium leprae.* BMC Genomics **10**:397. doi:10.1186/1471-2164-10-397.
33. **Newberg LA.** 2008. Significance of gapped sequence alignments. J. Comput. Biol. **15**:1187–1194.
34. **Hussain R, Shahid F, Zafar S, Dojki M, Dockrell HM.** 2004. Immune profiling of leprosy and tuberculosis patients to 15-mer peptides of *Mycobacterium leprae* and *M. tuberculosis* GroES in a BCG vaccinated area: implications for development of vaccine and diagnostic reagents. Immunology **111**:462–471.
35. **Maeda Y, Mukai T, Kai M, Fukutomi Y, Nomaguchi H, Abe C, Kobayashi K, Kitada S, Maekura R, Yano I, Ishii N, Mori T, Makino M.** 2007. Evaluation of major membrane protein-II as a tool for serodiagnosis of leprosy. FEMS Microbiol. Lett. **272**:202–205.
36. **Makino M, Maeda Y, Ishii N.** 2005. Immunostimulatory activity of major membrane protein-II from *Mycobacterium leprae.* Cell. Immunol. **233**:53–60.
37. **Spencer JS, Kim HJ, Marques AM, Gonzalez-Juarerro M, Lima MC, Vissa VD, Truman RW, Gennaro ML, Cho SN, Cole ST, Brennan PJ.** 2004. Comparative analysis of B- and T-cell epitopes of *Mycobacterium leprae* and *Mycobacterium tuberculosis* culture filtrate protein 10. Infect. Immun. **72**:3161–3170.
38. **Triccas JA, Roche PW, Winter N, Feng CG, Butlin CR, Britton WJ.** 1996. A 35-kilodalton protein is a major target of the human immune response to *Mycobacterium leprae.* Infect. Immun. **64**:5171–5177.
39. **de Souza GA, Softeland T, Koehler CJ, Thiede B, Wiker HG.** 2009. Validating divergent ORF annotation of the *Mycobacterium leprae* genome through a full translation data set and peptide identification by tandem mass spectrometry. Proteomics **9**:3233–3243.
40. **Petersen TN, Brunak S, von Heijne G, Nielsen H.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods **8**:785–786.
41. **Vordermeier HM, Hewinson RG, Wilkinson RJ, Wilkinson KA, Gideon HP, Young DB, Sampson SL.** 2012. Conserved immune recognition hierarchy of mycobacterial PE/PPE proteins during infection in natural hosts. PLoS One **7**:e40890. doi:10.1371/journal.pone.0040890.
42. **Nair S, Ramaswamy PA, Ghosh S, Joshi DC, Pathak N, Siddiqui I, Sharma PP, Hasnain SE, Mande SC, Mukhopadhyay T.** 2009. The PPE18 of *Mycobacterium tuberculosis* interacts with TLR2 and activates IL-10 induction in macrophage. J. Immunol. **183**:6269–6281.
43. **Wang P, Sidney J, Dow C, Mothe B, Sette A, Peters B.** 2008. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS Comput. Biol. **4**:e1000048. doi:10.1371/journal.pcbi.1000048.