Use of Protein Microarrays To Define the Humoral Immune Response in Leprosy Patients and Identification of Disease-State-Specific Antigenic Profiles[∇][†]

Nathan A. Groathouse,[1] Amol Amin,[1] Maria Angela M. Marques,[1] John S. Spencer,[1] Robert Gelber,[3] Dennis L. Knudson,[2] John T. Belisle,[1] Patrick J. Brennan,[1] and Richard A. Slayden[1]*

Department of Microbiology, Immunology, and Pathology[1] and Department of Bioagricultural Science and Pest Management,[2] Colorado State University, Fort Collins, Colorado 80523-1682, and Leonard Wood Memorial Leprosy Research Center, Cebu, Philippines[3]

Although the global prevalence of leprosy has decreased over the last few decades due to an effective multidrug regimen, large numbers of new cases are still being reported, raising questions as to the ability to identify patients likely to spread disease and the effects of chemotherapy on the overall incidence of leprosy. This can partially be attributed to the lack of diagnostic markers for different clinical states of the disease and the consequent implementation of differential, optimal drug therapeutic strategies. Accordingly, comparative bioinformatics and *Mycobacterium leprae* protein microarrays were applied to investigate whether leprosy patients with different clinical forms of the disease can be categorized based on differential humoral immune response patterns. Evaluation of sera from 20 clinically diagnosed leprosy patients using native protein and recombinant protein microarrays revealed unique disease-specific, humoral reactivity patterns. Statistical analysis of the serological patterns yielded distinct groups that correlated with phenolic glycolipid I reactivity and clinical diagnosis, thus demonstrating that leprosy patients, including those diagnosed with the paucibacillary, tuberculoid form of disease, can be classified based on humoral reactivity to a subset of *M. leprae* protein antigens produced in recombinant form.

Global leprosy disease prevalence has been drastically reduced, due largely to a World Health Organization-sponsored multidrug therapy elimination campaign (42). Incidence, as estimated by new case detection, however, remains high. Moreover, disease management and prevention in this new era of lowered prevalence have been hindered by the absence of tools that allow the objective diagnosis of disease and disease states, therefore providing a guide to preventative therapy and overall disease management. The identification of specific informative diagnostic antigens is one of the most difficult aspects in developing new diagnostic tools, and this is particularly true with leprosy, because there is a paucity of information involving the roles of many of the expressed proteins or the metabolic state of the organism throughout infection and disease progression.

The availability of the complete genome sequence and annotated coding capacity of *Mycobacterium leprae* provides a wealth of information that can be exploited for diagnostic purposes (4, 18). Of course, prospective antigens that may be relevant to disease diagnosis must then be validated experimentally. The major protein antigens of *M. leprae* were identified through subcellular fractionation of armadillo-derived *M. leprae* whole cells (16, 17, 21, 22, 27, 33, 34, 37). Recombinant forms of some of the more significant native proteins were subsequently created and tested (22, 27, 37). Recently, several groups have also used a postgenomic approach to discover new antigens for leprosy diagnosis (1, 2, 28, 36, 37). These studies all exploited genomic sequence for the identification of *M. leprae*-specific proteins or peptides that may be suitable for serodiagnosis of different disease states of leprosy. While many of these studies described novel antigens that show marked humoral and cellular immunogenicity, none employed protein-based microarrays.

The presence of antibodies follows an initial infection and precedes disease manifestations, allowing targeted chemoprophylaxis during the typical long incubation period (~5 years) of leprosy. Similar to the diagnosis of tuberculosis, where early detection of exposure and prompt chemoprophylaxis prevent the progression of disease, household contacts of multibacillary (MB) leprosy patients and exposed individuals would also benefit from early detection (10). Indeed, studies have shown that contacts of MB leprosy patients have an increased risk of developing leprosy themselves (41). It has also been found that contacts who have an antibody response to the *M. leprae*-specific phenolic glycolipid (phenolic glycolipid I [PGL-I]) have a much greater chance to develop clinical leprosy than those without an antibody response (3, 7, 19, 23). Yet almost half of those who have antibodies to PGL-I never develop leprosy, and half of those who develop leprosy never have PGL-I antibody. Thus, additional alternative markers have the promise of producing a predictive serodiagnostic tool.

Protein-based microarrays provide a consistent platform for studying humoral immune responses of a diverse group of patients to a wide variety of antigens for various infectious diseases in a high-throughput fashion (6, 9). In the present work, the

---

humoral immune response patterns of sera from patients clinically diagnosed with tuberculoid or lepromatous forms of leprosy (30) were evaluated with protein microarrays to define protein profiles reflective of specific disease states. The arrays were constructed either with proteins isolated from the cell wall and membrane of *M. leprae* or with a subset of recombinant proteins that are unique to *M. leprae* or have significant selectivity to *M. leprae*, according to stringent bioinformatics analysis. The results indicate that screening disease-state sera against protein-based microarrays can discern reactive antigens and patterns that are specific to disease classification. This work provides a foundation for the identification of novel diagnostic antigens relevant to the various clinical forms of leprosy, particularly tuberculoid.

MATERIALS AND METHODS

***M. leprae* patient serum samples.** Ten each of paucibacilliary (PB) and MB leprosy patients were diagnosed by clinical and histopathological criteria at the Leonard Wood Memorial Center for Leprosy Research, Cebu, Philippines. Leprosy was classified based on the Ridley-Jopling scheme by bacterial, histological, and clinical observation (30) carried out by experienced leprologists and a leprosy pathologist; no nerve biopsies were performed on the patients in this study. All sera were collected at the time of initial diagnosis before any antimicrobial therapy. Individuals clinically diagnosed with the lepromatous (LL) or borderline lepromatous (BL) forms of leprosy (samples L1 to L26) had an enzyme-linked immunosorbent assay (ELISA) value (optical density at 490 nm [OD_{490}]) against PGL-I of *M. leprae* (15) of 2.35 ± 0.28 and a mean bacterial index (BI) of 4.03 ± 0.62. Individuals clinically diagnosed with the tuberculoid (TT) or borderline tuberculoid (BT) forms of leprosy (samples T51 to T60) had an ELISA PGL-I value (OD_{490}) of 0.80 ± 0.36 and a mean BI of 0.48 ± 0.50. Details of the treatment of patients and clinical outcomes are presented in Table S1 in the supplemental material. Naive individuals from a site to which leprosy is not endemic (Colorado) provided control sera with an ELISA PGL-I value (OD_{490}) of 0.29 ± 0.03.

Isolation and purification of *M. leprae* subcellular fractions. Approximately 200 mg of *M. leprae* whole cells were purified from armadillo spleens and livers according to the Draper 3/77 protocol (33). Subcellular fractionation of *M. leprae* whole cells was achieved by sonic disruption (MSE Soniprep 150, MSE-Sonyo; Integrated Services, Palisades Park, NJ) for 30 cycles (60-s bursts followed by 60 s of cooling) in buffer consisting of 10 mM NH_4HCO_3 and 1 mM phenylmethylsulfonyl fluoride. The whole-cell sonicate was digested with 10 mg/ml of DNase and RNase for 1 h at 37°C (11). The pellet resulting from centrifugation at $27,000 \times g$ for 30 min provided the cell wall fraction, and the supernatant from this step was recentrifuged at $100,000 \times g$ for 2 h, yielding a second pellet of cytoplasmic membrane.

Final separation of cell wall and cytoplasmic membrane-associated proteins was achieved by electrophoresis on a preparative 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel performed under reducing conditions (20). On completion of electrophoresis, the gel was soaked in 20 mM NH_4HCO_3 for 30 min, followed by electrophoretic elution of the proteins using a Bio-Rad whole-gel elutor (Bio-Rad, Hercules, CA) at 250 mA for 2 h. The resulting protein fractions were frozen and lyophilized, resuspended in a 400-μl volume of sterile endotoxin-free water, and analyzed for content and purity by SDS-PAGE and silver staining (24). A periodic acid step was also incorporated to gauge the presence of or ensure the absence of lipoarabinomannan (40).

ELISA and Western blotting. High-affinity polystyrene microtiter plates (Immulon 4 HBX plates; Dynax, Alexandria, VA) were coated with protein antigens overnight at 4°C at concentrations ranging from 50 to 250 ng in 50 μl of phosphate-buffered saline (PBS) per well for purified antigens, up to 4 μg per well for membrane and cell wall fractions, and 0.5 to 2 μg per well for size-fractionated protein antigens. Wells were blocked with PBS containing 1% bovine serum albumin (Intergen Co., Purchase, NY) and 0.05% Tween 80 (Sigma Chemical Co., St. Louis, MO) for 1 h at room temperature. Polyclonal mouse sera and monoclonal antibodies were incubated at optimal dilutions in blocking buffer, as previously described (37). Unbound antibody was removed with PBS containing 0.05% Tween 80 without bovine serum albumin, and the secondary antibody conjugated to alkaline phosphatase (Sigma) was added and incubated for 2 h. Alkaline phosphatase activity was detected by the addition of a *p*-nitrophenylphosphate substrate. Western blots were prepared by transferring antigens run on SDS-PAGE (10% or 15% polyacrylamide gels) to a nitrocellulose membrane (39) and incubated with an antigen-specific primary antibody (37). Antibodies against *M. leprae* antigens were generated as described elsewhere (12, 16, 17,

22, 27, 34, 37); these are available from the Leprosy Research Support and Maintenance of an Armadillo Colony Post-Genome Era, Part I: Leprosy Research Support Contract (N01 AI-25469) at Colorado State University (5).

Comparative genomic and bioinformatics analysis. A global *in silico* identification of *t*argets-CROSS_MATCH (GISIT-cm)-approach was used to identify proteins that might be potential targets for further study. The GISIT-cm approach identifies unique proteins by comparing the *M. leprae* genome (GenBank entry NC_002677.fna) against other bacterial genomes. This was performed with the *Mycobacterium tuberculosis* genome (GenBank entry NC_000962.fna) by dividing it into 491 10-kb fragments, where each fragment contained a 1 kb-overlap with the previous fragment, using SPLITTER (EMBOSS package) (29). The data set of *M. tuberculosis* overlapping sequences was then used as the source for the masking sequence against the *M. leprae* genome using CROSS_MATCH (version 0.990329), which used a restricted Smith-Waterman (35) algorithm (13). CROSS_MATCH was run with a min-match value of 12 and a min-score value of 20, resulting in a masked *M. leprae* genome file where the sequences similar to those of *M. tuberculosis* were identified. ARTEMIS was then used to identify the open reading frames (ORFs) that were masked and to produce a masked data set of *M. leprae* ORFs. The *M. leprae* data set was opened in ARTEMIS (31), ORFs were selected, and a separate feature table of the 1,605 selected ORFs was prepared; this selection did not contain pseudogenes. Shell script and PERL scripts that read the FASTA (26) formatted file of masked proteins were written, producing a new file of proteins where each protein did not contain more than a specified percentage of cross-identity in the amino acid sequence. A value of 50% was used as the cutoff in this study. Uniqueness of identified ORFs to *M. leprae* was confirmed by BLASTN and BLASTP analysis against GenBank entries. The complete list of proteins identified using the GISIT-cm approach is in Table S2 in the supplemental material.

Production of recombinant proteins. Relevant genes were PCR amplified from *M. leprae* genomic DNA using Vent PFU DNA polymerase (Sigma). Primers for each gene were engineered to introduce NdeI and HindIII restriction enzyme sites into the 5′ and 3′ ends of the amplicon to facilitate direct cloning into the expression vector pET28(+) (EMD Biosciences, Inc., San Diego, CA). Each recombinant clone was verified by DNA sequencing. Recombinant protein production was achieved by introduction of the expression plasmid into *Escherichia coli* strain BL21(DE3) (Invitrogen Corp., Carlsbad, CA) by transformation and induction using the T7 polymerase with 0.3 mM isopropyl-β-D-thiogalactopyranoside. Recombinant proteins were released from *E. coli* by sonic disruption in buffer (Tris-HCl, pH 8.0, 2 μg/ml aprotinin, 1 μg/ml leupeptin, 1 μg/ml pepstatin, and 1 mM phenylmethylsulfonyl fluoride). The bacterial lysate was cleared by centrifugation, and the resulting supernatant was applied to an immobilized nickel-affinity column. Purified recombinant proteins were recovered from the affinity column with 50 mM imidazole and passed over a Detoxi-gel column (Pierce Biotechnology, Inc., Rockford, IL) to remove any contaminating endotoxin. Purity of recombinant proteins was assessed by SDS-PAGE, followed by silver staining. The final protein concentration was determined using the bicinchoninic assay (Pierce Biotechnology, Inc., Rockford, IL), and lipopolysaccharide contamination was evaluated by the Limulus amoebocyte lysate assay (Cambrex Corp., East Rutherford, N.J.).

Fabrication and immunoblotting of protein microarrays. *M. leprae* protein arrays were fabricated on glass slides with a 14 μM nitrocellulose film (FAST glass slides; Schleicher & Schuell BioScience, Inc., Keene, N.H.) using a Versarray Chipwriter Pro (Bio-Rad). Proteins fractions and buffer controls were printed in triplicate at approximately 0.2-mg/ml concentrations. Protein arrays were blocked for 1 h in protein array-blocking buffer (Schleicher & Schuell BioScience, Inc., Keene, NH) and incubated with serum (diluted 1:50) from patients or controls (primary antibody) at room temperature for 2 h. Visualization of primary antibody (Ab) (Sigma-Aldrich, St. Louis, MO) was achieved by incubation with Cy5- or Cy3-conjugated antihuman secondary Abs and scanning with a VersaArray ChipReader Pro (Bio-Rad). Fluorescence intensities were quantified using Spotfinder software (32, 38).

Array data analysis. Fluorescence intensities derived from each of the independent triplicate arrays were averaged to represent the response of each patient's serum sample. The resulting averaged intensities were then globally normalized for direct comparisons. Fluorescence intensities for each protein spot resulting from blotting with control serum were used to calculate the level of fluorescence intensity relative to background reactivity for each protein spot. The reactive index for each protein spot was calculated as the number of standard deviations relative to the average fluorescence intensity of all the spots. This statistical approach allowed for identification of protein antigens that were found to have significantly greater than average background reactivity. Hierarchal clustering and self-organizing map (SOM) analysis was performed on the entire data
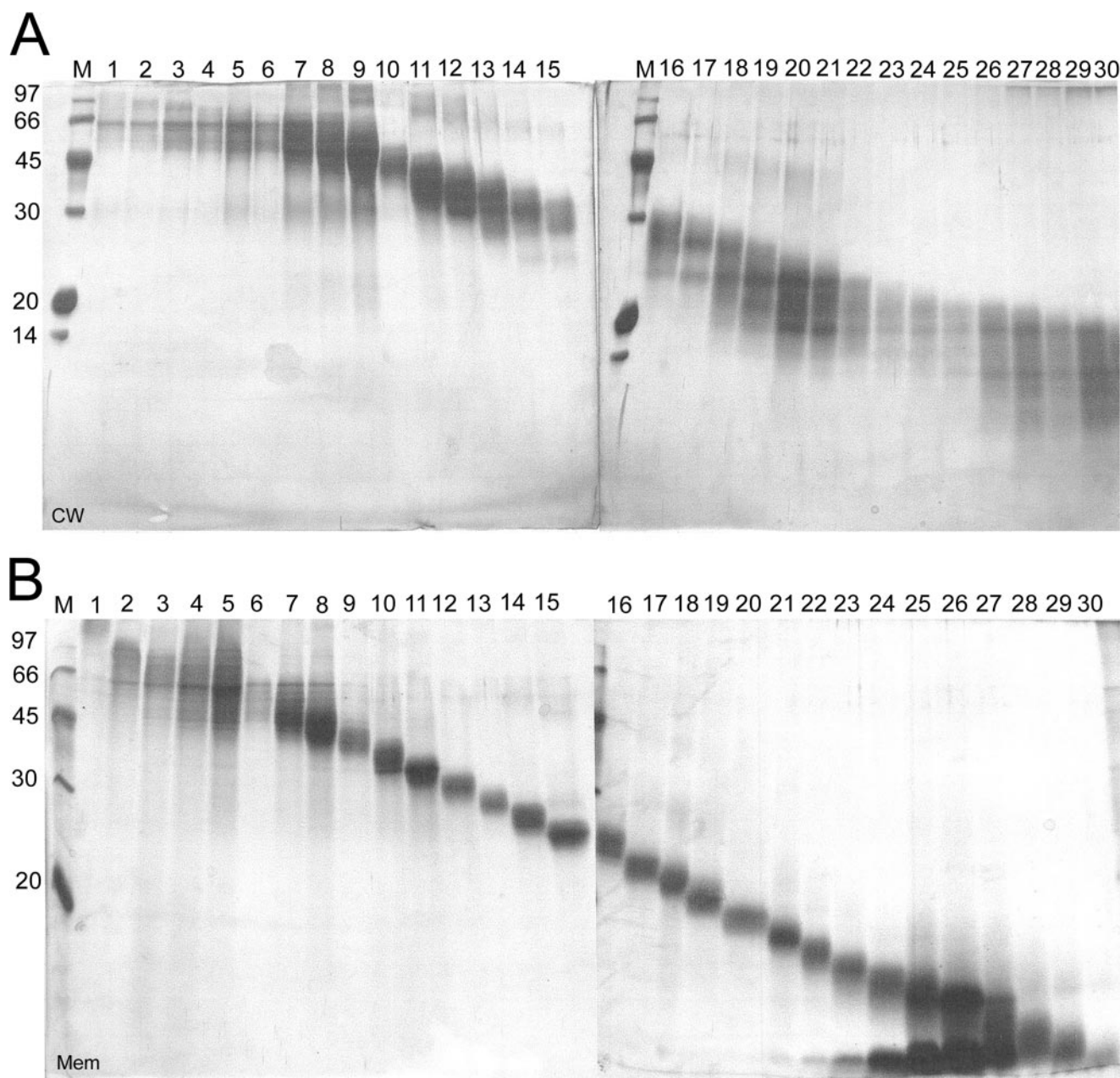
FIG. 1. SDS-PAGE gel migration analysis of the *M. leprae* native protein fractions used in the fabrication of the protein microarray. (A) Cell wall protein fractions and (B) membrane proteins fractions separated by electrophoretic elution and visualized by silver staining.

set (43); SOM is an unsupervised neural network model that effectively categorizes and clusters based on similarities in the antibody reactivity among groups.

RESULTS

Analysis of the humoral immune response using native-based protein arrays. Native protein arrays were printed with protein fractions derived from the *M. leprae* membrane and cell wall. These fractions were visualized by SDS-PAGE to evaluate overall sample fractionation and protein distribution (Fig. 1). Although the molecular weight range of proteins in each fraction was relatively narrow, previous quantitative analysis of two-dimensional gel patterns revealed that each protein

fraction used for array fabrication contained multiple proteins (21). To evaluate the potential distribution of a single protein among different protein fractions, Western blot analysis was performed (data not shown), demonstrating that known protein antigens were electrophoretically eluted into peak fractions with some overlap to adjacent fractions.

Native protein arrays were probed with serum obtained from patients clinically diagnosed with lepromatous or tuberculoid forms of leprosy (30). Immunologically naive individuals lack reactivity against any of the proteins on the array, but sera from individuals diagnosed with leprosy had different reactivity patterns. The reactive index for each protein antigen fraction
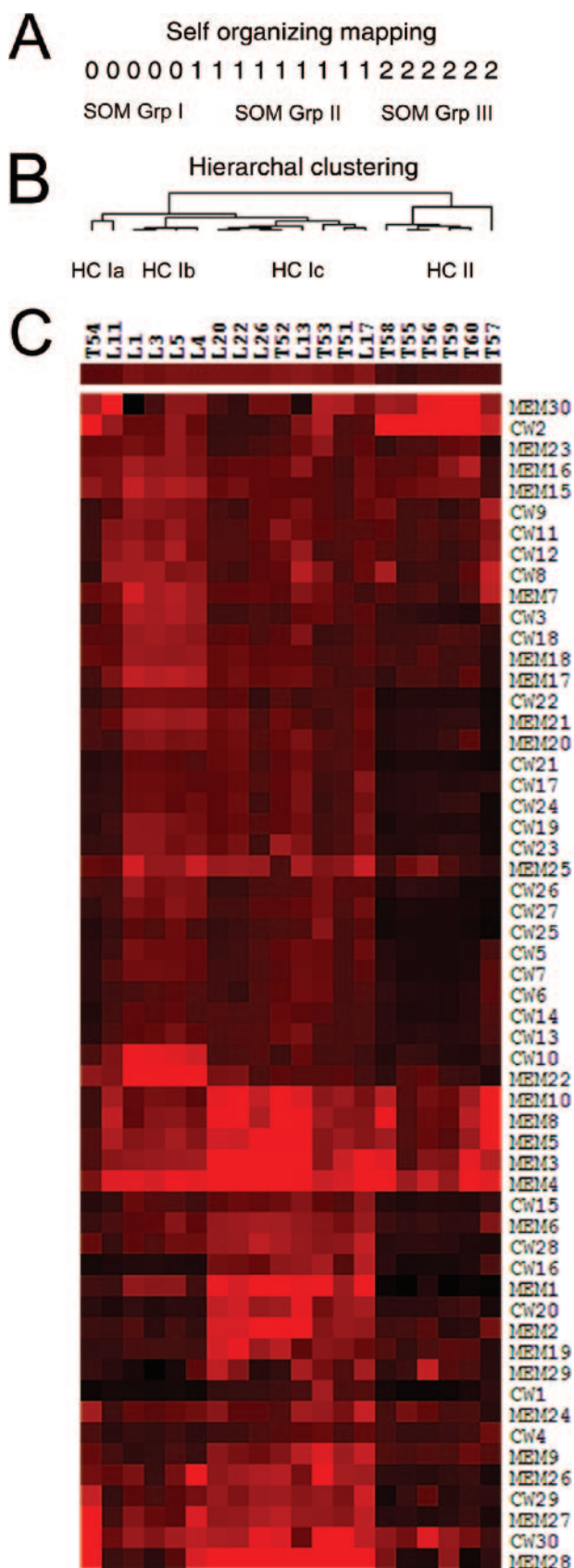
FIG. 2. Statistical analysis of humoral response patterns derived from *M. leprae* native protein microarrays. (A) Self-organizing map-

on the arrays was calculated and subjected to analysis to determine whether there were unique disease-specific patterns that correlate to disease diagnosis (Fig. 2C) (see Table S3 in the supplemental material). SOM analysis of the reactive index for each protein fraction assigned patients into three groups based on the reactive patterns of their sera (Fig. 2A). SOM group I was predominately composed of patients that had been clinically diagnosed with the lepromatous form of leprosy (SOM 0; $n = 4/5$). SOM group III was entirely composed of patients clinically diagnosed with the tuberculoid form of leprosy (SOM 2; $n = 6/6$). The largest and most clinically diverse group was SOM group II, which had three patients clinically diagnosed with the tuberculoid form of leprosy (SOM 1; $n = 3/9$) and six patients clinically diagnosed with the lepromatous form of leprosy (SOM 1; $n = 6/9$). While there were some protein fractions recognized in common by all patient sera, many of the protein fractions were uniquely recognized by sera from patients assigned to a single SOM group (Fig. 3A to C; Table 1). Specific reactivity patterns which correlate with different clinical states of disease were seen using protein microarrays and statistical analysis.

Hierarchal clustering analysis was also performed on the data set. This statistical approach organized patients into two rather than three major groups (Fig. 2B). The main difference was that patients assigned to SOM groups I and II were combined, and several smaller subdivisions, namely HC groups Ia to Ic, emerged. Overall, this analysis organized the majority of the lepromatous patients together (HC groups Ib and Ic; $n = 9/10$) with a small group (HC group Ia) as statistical outliers with one lepromatous and one tuberculoid patient. Importantly, the second major group, revealed by hierarchal clustering analysis (HC group II), contained the same patients as SOM group III, which was wholly comprised of patients that were clinically diagnosed with the TT form of leprosy.

Further examination of the analysis clearly revealed a group of patients assigned to SOM group II or HC group Ib by SOM or hierarchical clustering analysis, respectively, whose sera had similar reactivity patterns despite clinical diagnosis and that were different from other lepromatous and tuberculoid patients, favoring classification in a more intermediary position within the Ridley-Jopling clinical spectrum (30). These observations support the case for a borderline form of disease (BT, BB, or BL). Statistical analyses supported the case for the existence of unique patterns of serological reactivity to *M. leprae* protein fractions for different clinical states of disease, thus substantiating this approach of using *M. leprae* protein microarrays for the identification of disease state-specific reactive patterns, particularly for the tuberculoid from of disease.

The complexity of the native protein fractions hindered precise identification of all of the potentially reactive proteins within each fraction by mass spectrometry, N-terminal sequencing, or Western blotting. Since certain dominant protein antigens were known

ping and (B) hierarchal cluster analysis of reactive indices from patient sera on native protein microarrays. (C) Map of serum reactivity patterns for each native protein fraction. The reactive indices for each protein were calculated and the statistical analysis performed as described in Materials and Methods.
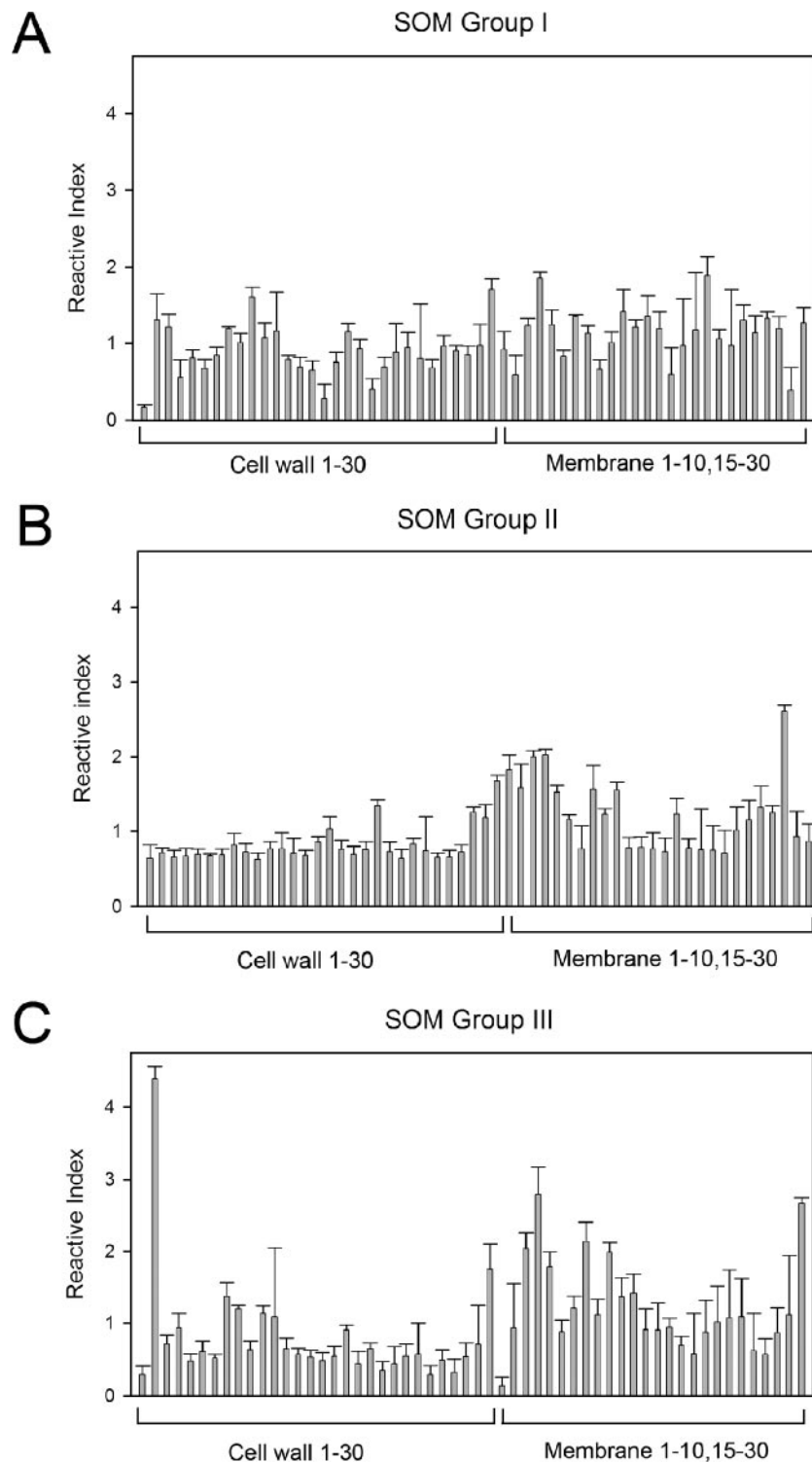
FIG. 3. Reactive profiles of patient groups assigned by self-organizing mapping. SOM group I (A), group II (B), and group III (C) as determined from native microarray analysis. The reactive indices for each protein were calculated as described in Materials and Methods.

to be present in the spotted native protein fractions as determined by application of antigen-specific monoclonal antibodies, patient sera reactive to these spots implicated reaction to these precise proteins (see Table S4 in the supplemental material). Further-

more, sera reactive to multiple fractions containing a common protein strongly implicated a particular antigen as the immuno-dominant protein in those fractions. However, this is not a defin-itive identification system. Accordingly, recombinant protein-

TABLE 1. Dominant protein fractions for each SOM group identified by patient sera on native protein microarrays

Protein fraction[c]	Mean RI[a,b]		
	SOM I	SOM II	SOM III
CW10	**1.60 ± 0.43**	0.63 ± 0.09	0.64 ± 0.21
CW2	**1.31 ± 0.67**	0.71 ± 0.30	**4.40 ± 1.33**
CW20	0.40 ± 0.07	**1.34 ± 0.25**	0.65 ± 0.20
CW30	**1.70 ± 1.06**	**1.67 ± 0.60**	**1.75 ± 0.65**
CW8	1.19 ± 0.38	0.81 ± 0.24	**1.38 ± 0.65**
MEM1	0.92 ± 0.27	**1.83 ± 0.27**	0.14 ± 0.18
MEM10	1.01 ± 0.37	**1.55 ± 0.42**	**1.98 ± 0.85**
MEM15	**1.42 ± 0.12**	0.78 ± 0.10	1.37 ± 0.29
MEM16	1.22 ± 0.11	0.79 ± 0.11	**1.42 ± 0.48**
MEM17	**1.36 ± 0.28**	0.77 ± 0.11	0.91 ± 0.28
MEM2	0.59 ± 0.12	**1.58 ± 0.39**	0.94 ± 0.33
MEM22	**1.89 ± 0.36**	0.76 ± 0.16	0.88 ± 0.34
MEM25	**1.31 ± 0.28**	1.16 ± 0.26	1.09 ± 0.47
MEM27	**1.33 ± 0.78**	1.25 ± 0.14	0.58 ± 0.16
MEM28	1.19 ± 0.95	**2.61 ± 0.83**	0.88 ± 0.68
MEM3	1.24 ± 0.18	**2.00 ± 0.44**	**2.04 ± 0.62**
MEM30	1.27 ± 1.01	0.87 ± 0.43	**2.67 ± 0.80**
MEM4	**1.85 ± 0.32**	**2.03 ± 0.32**	**2.79 ± 0.67**
MEM5	1.24 ± 0.32	**1.52 ± 0.34**	**1.78 ± 0.72**
MEM7	**1.36 ± 0.20**	0.77 ± 0.13	1.21 ± 0.52
MEM8	1.13 ± 0.46	**1.56 ± 0.39**	**2.14 ± 1.04**

[a] Mean reactive index (RI) is defined as the number of standard deviations of a normalized fluorescence intensity above the background level.
[b] SOM group (SOM I to III) assignments are from statistical analysis of immunoreactivity on native protein microarrays probed with sera from patients clinically diagnosed with the TT or LL form of disease.
[c] Dominant protein fractions for each SOM group. The top 10 immunoreactive protein fractions are denoted by boldfaced values for each SOM group (SOM I to III).

TABLE 2. Dominant protein fractions for each SOM group identified by patient sera on recombinant protein microarrays

ORF	Protein[b]	Mean RI[a]		
		SOM I	SOM II	SOM III
ml0008	Grp I Ag-1	**1.71 ± 0.51**	1.00 ± 0.24	0.84 ± 0.43
ml0957	Grp I Ag-2	**2.09 ± 0.51**	1.01 ± 0.27	0.57 ± 0.31
ml1877	Grp II Ag-1/Tuf	0.32 ± 0.08	**1.22 ± 0.24**	0.49 ± 0.20
ml1829	Grp II Ag-2	0.86 ± 0.37	**1.37 ± 0.24**	0.33 ± 0.15
ml0126	Grp II Ag-3	0.34 ± 0.13	**1.11 ± 0.22**	0.23 ± 0.12
ml0396	Grp II Ag-4	0.61 ± 0.19	**1.25 ± 0.48**	0.23 ± 0.09
ml1419	Grp III Ag-1	0.41 ± 0.13	1.03 ± 0.17	**1.45 ± 0.63**
ml1057	Grp III Ag-2	0.27 ± 0.16	0.18 ± 0.09	**1.04 ± 0.81**
ml1915	Un Ag-1	<u>**1.37 ± 0.29**</u>	<u>**1.42 ± 0.37**</u>	<u>**1.54 ± 0.69**</u>
ml0050	Un Ag-2/CFP-10	<u>**6.19 ± 1.91**</u>	<u>**1.71 ± 0.69**</u>	<u>**6.99 ± 3.41**</u>

[a] Mean reactive index (RI) is defined as the number of standard deviations of a normalized fluorescence intensity above the background level. SOM group (SOM I to III) assignments are from statistical analysis of immunoreactivity on native protein microarrays probed with sera from patients clinically diagnosed with the TT or LL form of disease.
[b] Proteins were designated universal (Un) or SOM Grp I, Grp II, or Grp III based on patient serum reactivity and SOM analysis from Fig. 2. Significant RIs are shown in boldface. Underlined values are RI values for multiple SOM groups.

based arrays were applied towards more definitive identification of antigenic proteins.

Analysis of humoral immune response to selected recombinant proteins. GISIT-cm was performed to identify proteins unique to *M. leprae* compared to other mycobacterial species. Of 1,605 *M. leprae*-encoded proteins, 214 were found to have 50% or greater selectivity to *M. leprae* among contiguous protein sequence while 160 were considered 100% unique (see Table S2 in the supplemental material). *M. leprae*-unique proteins identified from our bioinformatic approach are in agreement with those recently identified by others (2). Notably, only 6 of the 160 unique proteins were annotated to a putative function, whereas the others were annotated as hypothetical proteins (18). Using information obtained through bioinformatic analyses and reactivity on native protein arrays, 18 proteins were selected for recombinant production and purification.

This set of recombinant proteins was evaluated through microarray technology to evaluate and identify a subset of proteins that can serve as leprosy-specific disease state antigens. Upon screening with patient sera, the reactive index for each recombinant protein was calculated and subjected to SOM analysis (Table 2). Similar to what was observed using native arrays, antigenic proteins fell into three basic diagnostic categories: those recognized by tuberculoid patients, those recognized by lepromatous patients, and those recognized by a subset of tuberculoid and lepromatous patients. The last group may represent borderline forms of disease. Group I antigen I (Grp I Ag-1) (ML0008) and Grp I Ag-2 (ML0957) were recognized only by sera of patients clinically diagnosed with the

lepromatous form of the disease. Grp II Ag-1/Tuf (ML1877), Grp II Ag-2 (ML1829), Grp II Ag-3 (ML0126), and Grp II Ag-4 (ML0396) were identified as being differentially recognized by sera of patients thought to have an intermediate form of leprosy based on statistical analysis of reactivity patterns using protein microarrays. Grp III Ag-1 (ML1419) and Grp III Ag-2 (ML1057) were recognized by sera from patients that were clinically diagnosed with the tuberculoid form of disease. Interestingly, a recent study did not find ML1057 to elicit significant gamma interferon (IFN-γ) production in leprosy patients (2). Two recombinant proteins were found to be recognized universally in this study for all patients infected with leprosy. Accordingly, these proteins, universal Ag-1 (ML1915) and universal Ag-2/CFP-10 (ML0050), are predictive antigens diagnostic for exposure to and infection with *M. leprae*. Others have reported immunoreactivity of sera of leprosy patients to ML1877 and ML0050, substantiating the discovery of these proteins as seroreactive antigens, and the use of ML0050 as a universal antigen for exposure to *M. leprae* (2, 28, 37). Taken together, these results demonstrate that as few as 10 recombinant proteins can be used to acquire meaningful information about a disease state and that directed recombinant arrays can yield information equivalent to what was provided by full native protein microarrays but with the added benefit of using proteins with known identity.

Correlation of protein array classification and reactivity to PGL-I. A current serodiagnostic test that is able to identify patients with *M. leprae* infection is based on *M. leprae*-specific PGL-I. This antigen has been shown to be a marker for bacterial load, with antibody levels correlating with the spectrum of disease (10, 14, 25). Accordingly, sera from all the patients used in this study were evaluated by ELISA for PGL-I seroreactivity (Table 3). Overall, patients diagnosed with the PB form of disease had lower ELISA PGL-I values (OD$_{490}$) (0.80 ± 0.36) and patients diagnosed with the MB form of disease had greater PGL-I values (2.35 ± 0.28), which is partially concordant with immunoreactivity patterning (Fig. 4). As discussed previously, patients were categorized into three

TABLE 3. Clinical backgrounds of leprosy patients used in this study

Patient no.	Age	Sex[a]	Duration of illness prior to diagnosis	Treatment status at time of collection	Clinical diagnosis	Histological diagnosis	Preliminary skin smear Avg BI[b]	Preliminary skin smear Site BI[c]	Histological BI[d]	LEPRA reaction	PGL-1 ELISA reactivity (OD$_{490}$)
L1	27	M	8 mo	Untreated	BL	BL	4.5+	4+	5+	None	1.88
L3	30	M	10 yr	Untreated	BL	BL	3.5+	4+	6+	None	2.47
L4	21	M	2 yr	Untreated	LL	LL	5+	4+	6+	None	2.17
L5	27	M	1 yr	Untreated	LL	LL	4+	4+	6+	None	2.02
L11	18	M	3 yr	Untreated	LL	BL	4.16+	4+	5+	None	2.91
L13	36	M	4 yr	Untreated	BL	BL	3+	4+	5+	None	2.47
L17	29	M	2 yr	Untreated	LL	LL	4.83+	5+	6+	None	2.33
L20	33	M	1 yr	Untreated	BL	BL	3.5+	4+	5+	None	1.96
L22	44	M	3 yr	Untreated	LL	LL	4+	4+	6+	None	2.42
L26	16	M	4 yr	Untreated	LL	LL	3.83+	4+	5+	None	2.9
T51	32	F	2 yr	Untreated	TT	TT	0	0	0	None	0.67
T52	22	M	1 yr	Untreated	BT	BT	0.5+	0	0	None	0.79
T53	55	M	1 yr	Untreated	BL	BT	0.66+	1+	0	None	0.87
T54											0.63
T55	23	M	1 mo	Untreated	BT	TT	0	0	0	None	1.01
T56	48	M	1 yr	Untreated	BL	BT	0	0	0	None	0.71
T57	51	M	10 yr	Untreated	BT	BT			0	None	0.63
T58	58	M	5 mo	Untreated	BL	BT	1.5+	2+	0	On reaction (RR+); No steroids taken	1.2
T59	41	M	6 mo	Untreated	BT	BT	0.7+	1+	0	None	0.31
T60	35	M	10 yr	Untreated	BT	BT	0.5+	1+	0	On reaction (RR+); No steroids taken	1.16

[a] M, male; F, female.
[b] Average bacterial index (BI) is the mean value for six smear sites.
[c] BI from a skin smear of the biopsy site.
[d] BI from the histological section of the biopsy.

groups based on serum reactivity. SOM group I consisted of PB patients, and SOM group III consisted of MB patients, whereas SOM group II contained both PB and MB patients. Although there was a general concordance between the clinical diagnosis of the patients and the statistical categorization, a different state of disease progression, perhaps borderline forms of disease, is indicated by the statistical elucidation of SOM group II. Importantly, patients grouped in SOM group II had a mean PGL-I ELISA value (OD$_{490}$) of 1.80 ± 0.76, which is a value that was in between the mean value for patients clinically diagnosed with PB and patients clinically diagnosed with MB forms of disease. The conclusion that individuals categorized in SOM group II are indeed at a different stage of disease progression is supported by this observation and is consistent with elevated antibodies against PGL-I being associated with spectrum of disease and relapse; in fact it has been suggested that PB leprosy patients with elevated antibodies should be treated as MB leprosy patients (10).

DISCUSSION

One of the most challenging tasks in developing disease state-specific serodiagnostics is the identification of discriminating antigens that differentiate between exposure and clinical stage of disease with high sensitivity and specificity. Screening sera from a large number of patients diagnosed with various states of disease against the entire leprosy proteome offers the potential for facile identification of such selective antigens. However, the resources to accomplish such an extensive enterprise with leprosy are not available. Therefore, to utilize microarray technology for the identification of novel diagnostic antigens, native proteins were obtained by subcellular fractionation of *M. leprae* and selected proteins were identified for recombinant antigen production based on bioinformatic analyses. Specifically, for selection of recombinant proteins, comparative analysis of the leprosy genome against those of closely related organisms was performed to identify gene products that are unique to *M. leprae*, with a consequent
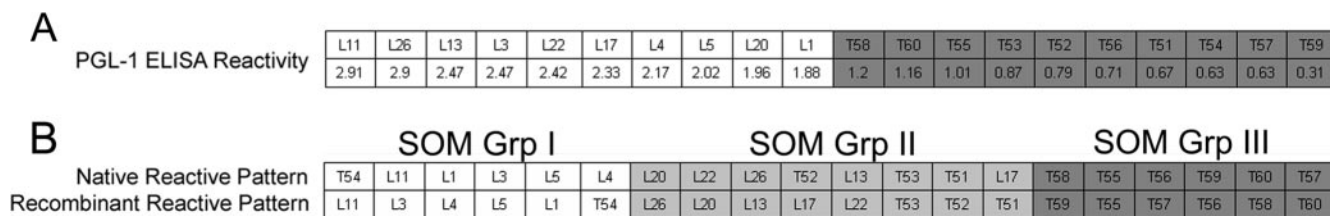

FIG. 4. Statistical analysis and categorization of disease state based on patient sera reactivity derived from native and recombinant-based protein microarrays. (A) The PGL-I ELISA reactivity correlated with the SOM analysis of reactivity patterns from (B) native and recombinant protein microarray analysis. The T series of patient sera and L series of patient sera are described in Materials and Methods. The shaded areas highlight the different disease states and statistical grouping.

high degree of serological specificity. Such an approach has been successfully used to identify highly specific antigens for tuberculosis diagnostics (8).

Currently the serodiagnosis of leprosy has been largely confined to the presence of immunoglobulin M antibodies to the *M. leprae*-specific PGL-I. Though antibodies to PGL-I are present in more than 90% of untreated MB lepromatous patients, only a limited number of patients at the PB/tuberculoid end of the disease spectrum are reactive (23, 25). Thus, the PB state, with low levels of circulating specific antibodies, absence of acid-fast bacilli, and clinical similarities to numerous other granulomatous processes, is difficult to diagnose (19). Adding to the complication of leprosy diagnosis is the requirement for highly trained clinicians that can differentiate clinical states of disease and categorize patients within the disease spectrum. In the absence of such experienced clinicians, diagnoses of each clinical form of leprosy is subjective (14). Depending on the categorization, the chemotherapeutic regimen varies: 6 months of multidrug therapy for tuberculoid patients compared to 12 months or more for lepromatous patients. Therefore, to enhance leprosy diagnosis and treatment, particularly for tuberculoid patients, an accurate diagnostic tool that provides a clear definition and a benchmark for disease progression is desirable.

In an attempt to improve diagnostics, multiple tests have been developed for leprosy; however, they lack either specificity or sensitivity for the detection of asymptomatic infections and disease progression. Recently studies employing bioinformatics and experimental approaches to evaluate individual *M. leprae* proteins or small sets of proteins as potential serodiagnostic or T-cell antigens have been performed (1, 2, 14, 28, 36). Reed and colleagues (28) identified 14 recombinant *M. leprae* proteins that strongly react to sera of LL patients, and two of these antigens (Ml0405 and Ml2331) demonstrated the ability to detect BL patients and, in combination, enhanced serological detection with PGL-I. Geluk et al. (14) also evaluated a relatively large number of recombinant *M. leprae* proteins for reactivity to T cells. This work demonstrated five antigens (Ml0576, Ml1989, Ml1990, Ml2283, and Ml2567) that induced significant IFN-γ levels in PB leprosy patients, reactional leprosy patients, and contacts but not in most MB patients or controls. Recently, recombinant proteins (Ml0008, Ml0126, Ml1057, and Ml2567) and 58 peptides were tested by us for IFN-γ responses in peripheral blood mononuclear cells from leprosy patients seeking epitopes that would increase specificity (36). The responses to the four recombinant proteins gave higher levels of IFN-γ production but less specificity than the peptides, with 35 of the peptides giving high responses only in the case of PB and household contacts. Another study evaluated the immunogenicity of 12 recombinant proteins by measuring the reactivity of circulating antibody and IFN-γ responses. Both humoral and cellular immunogenicity was observed for two antigens (Ml0308 and Ml2498) for PB and MB patients (2). It is interesting to note that there is limited overlap between the *M. leprae* proteins studied in previous work (36) and the 18 recombinant proteins evaluated in this study. However, the methods for selecting and screening of potential antigens in these studies were dramatically different. Overall, none of these studies identified unique antigens capable of distinguishing patients with PB versus MB forms of disease.

In our current studies, evaluation of serological reactivities for 20 patients clinically defined with either the PB or MB form of disease led to the identification of 10 proteins, allowing classification of patients into 3 categories. Sera from six PB patients uniquely recognized Ml0008 and Ml0957, and sera from six MB patients uniquely recognized, Ml1419 and Ml1057. Sera from the remaining PB and MB patients reacted with Ml1877, Ml1829, Ml0126, and Ml0396, giving rise to a third category. All patient sera had reactivity to Ml1915 and Ml0050, providing broad controls, similar to previous studies discussed. Identification of these 10 antigens based on serological activity established that a limited number of antigens can be used to categorize patients into groups consistent with clinical diagnosis based solely on nonsubjective criteria.

An interesting finding in this study is the statistical identification of a set of patients clinically diagnosed with either the PB or MB form of disease with similar humoral reactivity profiles (SOM group II). One possibility that may account for this is that these patients, with different clinical diagnoses, intermediate PGL-I reactivity, and different bacterial burdens, may be progressing along the clinical spectrum of disease. In such a case, the ability of these 10 antigens to distinguish true PB patients from those progressing towards the MB form of disease would have significant utility in leprosy control programs and in limiting the transmission of *M. leprae*. It has been reported that PB patients with weak PGL-I antibody responses are not associated with the spread of disease, whereas PB leprosy patients with elevated antibody responses transmit bacilli. Therefore, PB patients in this study that were categorized into SOM group II based on seroreactivity and that have elevated PGL-I reactivity might be progressing to the MB state. Fully realizing the potential of the antigens described in this study will require a larger cohort of patients and follow-up studies on disease progression.

A second aspect of this work was the use of complex subcellular protein fractions from an obligate intracellular pathogen to fabricate microarrays seeking to define unique serological reactivity profiles. While precise antigen identifications were not made, the use of native protein microarrays proved useful for discerning unique patterns in leprosy patients. Since the native protein fractions were limiting, extensive antigen identification could not be performed. Nevertheless, it was interesting to note that regardless of whether native protein fractions or recombinant proteins were used, patients sera grouped equally well based on disease state. The data obtained with the native fractions also indicate that there are potentially more diagnostic antigens to be discovered. Protein array technology may not yet be applicable as a field diagnostic in regions of endemicity. It is, however, a powerful tool for antigen discovery and could be applied to other clinically relevant research questions, including the identification of serodiagnostic antigens that can be used to monitor the success or failure of therapy.

REFERENCES

1. **Araoz, R., N. Honore, S. Banu, C. Demangel, Y. Cissoko, C. Arama, M. K. Mafij Uddin, S. K. Abdul Hadi, M. Monot, S. N. Cho, B. Ji, P. J. Brennan, S. Sow, and S. T. Cole.** 2006. Towards an immunodiagnostic test for leprosy. Microbes Infect. [Epub ahead of print.]

2. **Araoz, R., N. Honore, S. Cho, J. P. Kim, S. N. Cho, M. Monot, C. Demangel, P. J. Brennan, and S. T. Cole.** 2006. Antigen discovery: a postgenomic approach to leprosy diagnosis. Infect. Immun. **74:**175–182.

3. **Cardona-Castro, N. M., S. Restrepo-Jaramillo, M. Gil de la Ossa, and P. J. Brennan.** 2005. Infection by Mycobacterium leprae of household contacts of lepromatous leprosy patients from a post-elimination leprosy region of Colombia. Mem. Inst. Oswaldo Cruz **100:**703–707.

4. **Cole, S. T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honore, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell.** 2001. Massive gene decay in the leprosy bacillus. Nature **409:**1007–1011.

5. **Colorado State University.** Leprosy research support. College of Veterinary Medicine and Biological Sciences, Colorado State University, Fort Collins, Colo. [Online.] http://www.cvmbs.colostate.edu/mip/leprosy/index.html.

6. **Davies, D. H., X. Liang, J. E. Hernandez, A. Randall, S. Hirst, Y. Mu, K. M. Romero, T. T. Nguyen, M. Kalantari-Dehaghi, S. Crotty, P. Baldi, L. P. Villarreal, and P. L. Felgner.** 2005. Profiling the humoral immune response to infection by using proteome microarrays: high-throughput vaccine and diagnostic antigen discovery. Proc. Natl. Acad. Sci. USA **102:**547–552.

7. **Desforges, S., P. Bobin, B. Brethes, M. Huerre, J. P. Moreau, and M. A. Bach.** 1989. Specific anti-M leprae PGL-I antibodies and Mitsuda reaction in the management of household contacts in New Caledonia. Int. J. Lepr. Other Mycobact. Dis. **57:**794–800.

8. **Dietrich, J., C. V. Lundberg, and P. Andersen.** 2006. TB vaccine strategies—what is needed to solve a complex problem? Tuberculosis (Edinburgh) **86:**163–168.

9. **Doolan, D. L., J. C. Aguiar, W. R. Weiss, A. Sette, P. L. Felgner, D. P. Regis, P. Quinones-Casas, J. R. Yates III, P. L. Blair, T. L. Richie, S. L. Hoffman, and D. J. Carucci.** 2003. Utilization of genomic sequence information to develop malaria vaccines. J. Exp. Biol. **206:**3789–3802.

10. **Douglas, J. T., R. V. Cellona, T. T. Fajardo, Jr., R. M. Abalos, M. V. Balagon, and P. R. Klatser.** 2004. Prospective study of serological conversion as a risk factor for development of leprosy among household contacts. Clin. Diagn. Lab Immunol. **11:**897–900.

11. **Ehrenberg, J. P., and N. Gebre.** 1987. Analysis of the antigenic profile of Mycobacterium leprae: cross-reactive and unique specificities of human and rabbit antibodies. Scand. J. Immunol. **26:**673–681.

12. **Engers, H. D., et al.** 1985. Results of a World Health Organization-sponsored workshop on monoclonal antibodies to *Mycobacterium leprae.* Infect. Immun. **48:**603–605.

13. **Ewing, B., L. Hillier, M. C. Wendl, and P. Green.** 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. **8:**175–185.

14. **Geluk, A., M. R. Klein, K. L. Franken, K. E. van Meijgaarden, B. Wieles, K. C. Pereira, S. Buhrer-Sekula, P. R. Klatser, P. J. Brennan, J. S. Spencer, D. L. Williams, M. C. Pessolani, E. P. Sampaio, and T. H. Ottenhoff.** 2005. Postgenomic approach to identify novel *Mycobacterium leprae* antigens with potential to improve immunodiagnosis of infection. Infect. Immun. **73:**5636–5644.

15. **Hunter, S. W., T. Fujiwara, and P. J. Brennan.** 1982. Structure and antigenicity of the major specific glycolipid antigen of Mycobacterium leprae. J. Biol. Chem. **257:**15072–15078.

16. **Hunter, S. W., M. McNeil, R. L. Modlin, V. Mehra, B. R. Bloom, and P. J. Brennan.** 1989. Isolation and characterization of the highly immunogenic cell wall-associated protein of *Mycobacterium leprae.* J. Immunol. **142:**2864–2872.

17. **Hunter, S. W., B. Rivoire, V. Mehra, B. R. Bloom, and P. J. Brennan.** 1990. The major native proteins of the leprosy bacillus. J. Biol. Chem. **265:**14065–14068.

18. **Institut Pasteur.** 2004. Leproma world-wide web server. [Online.] http://genolist.pasteur.fr/Leproma/.

19. **Jardim, M. R., S. L. Antunes, B. Simons, J. G. Wildenbeest, J. A. Nery, X. Illarramendi, M. O. Moraes, A. N. Martinez, L. Oskam, W. R. Faber, E. N. Sarno, E. P. Sampaio, and S. Buhrer-Sekula.** 2005. Role of PGL-I antibody detection in the diagnosis of pure neural leprosy. Lepr. Rev. **76:**232–240.

20. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.

21. **Marques, M. A., B. J. Espinosa, E. K. Xavier da Silveira, M. C. Pessolani, A. Chapeaurouge, J. Perales, K. M. Dobos, J. T. Belisle, J. S. Spencer, and P. J. Brennan.** 2004. Continued proteomic analysis of Mycobacterium leprae subcellular fractions. Proteomics **4:**2942–2953.

22. **Mehra, V., B. R. Bloom, A. C. Bajardi, C. L. Grisso, P. A. Sieling, D. Alland, J. Convit, X. D. Fan, S. W. Hunter, P. J. Brennan, et al.** 1992. A major T cell antigen of *Mycobacterium leprae* is a 10-kD heat-shock cognate protein. J. Exp. Med. **175:**275–284.

23. **Moet, F. J., A. Meima, L. Oskam, and J. H. Richardus.** 2004. Risk factors for the development of clinical leprosy among contacts, and their relevance for targeted interventions. Lepr. Rev. **75:**310–326.

24. **Morrissey, J. H.** 1981. Silver stain for proteins in polyacrylamide gels: a modified procedure with enhanced uniform sensitivity. Anal. Biochem. **117:**307–310.

25. **Oskam, L., E. Slim, and S. Buhrer-Sekula.** 2003. Serology: recent developments, strengths, limitations and prospects: a state of the art overview. Lepr. Rev. **74:**196–205.

26. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.

27. **Pessolani, M. C., D. R. Smith, B. Rivoire, J. McCormick, S. A. Hefta, S. T. Cole, and P. J. Brennan.** 1994. Purification, characterization, gene sequence, and significance of a bacterioferritin from *Mycobacterium leprae.* J. Exp. Med. **180:**319–327.

28. **Reece, S. T., G. Ireton, R. Mohamath, J. Guderian, W. Goto, R. Gelber, N. Groathouse, J. Spencer, P. Brennan, and S. G. Reed.** 2006. ML0405 and ML2331 are antigens of Mycobacterium leprae with potential for diagnosis of leprosy. Clin. Vaccine Immunol. **13:**333–340.

29. **Rice, P., I. Longden, and A. Bleasby.** 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. **16:**276–277.

30. **Ridley, D. S., and W. H. Jopling.** 1966. Classification of leprosy according to immunity. A five-group system. Int. J. Lepr. Other Mycobact. Dis. **34:**255–273.

31. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945.

32. **Saeed, A. I., V. Sharov, J. White, J. Li, W. Liang, N. Bhagabati, J. Braisted, M. Klapa, T. Currier, M. Thiagarajan, A. Sturn, M. Snuffin, A. Rezantsev, D. Popov, A. Ryltsov, E. Kostukovich, I. Borisovsky, Z. Liu, A. Vinsavich, V. Trush, and J. Quackenbush.** 2003. TM4: a free, open-source system for microarray data management and analysis. BioTechniques **34:**374–378.

33. **Shepard, C. C., P. Draper, R. J. Rees, and C. Lowe.** 1980. Effect of purification steps on the immunogenicity of *Mycobacterium leprae.* Br. J. Exp. Pathol. **61:**376–379.

34. **Silbaq, F. S., S. N. Cho, S. T. Cole, and P. J. Brennan.** 1998. Characterization of a 34-kilodalton protein of *Mycobacterium leprae* that is isologous to the immunodominant 34-kilodalton antigen of *Mycobacterium paratuberculosis.* Infect. Immun. **66:**5576–5579.

35. **Smith, T. F., and M. S. Waterman.** 1981. Identification of common molecular subsequences. J. Mol. Biol. **147:**195–197.

36. **Spencer, J. S., H. M. Dockrell, H. J. Kim, M. A. Marques, D. L. Williams, M. V. Martins, M. L. Martins, M. C. Lima, E. N. Sarno, G. M. Pereira, H. Matos, L. S. Fonseca, E. P. Sampaio, T. H. Ottenhoff, A. Geluk, S. N. Cho, N. G. Stoker, S. T. Cole, P. J. Brennan, and M. C. Pessolani.** 2005. Identification of specific proteins and peptides in mycobacterium leprae suitable for the selective diagnosis of leprosy. J. Immunol. **175:**7930–7938.

37. **Spencer, J. S., H. J. Kim, A. M. Marques, M. Gonzalez-Juarrero, M. C. Lima, V. D. Vissa, R. W. Truman, M. L. Gennaro, S. N. Cho, S. T. Cole, and P. J. Brennan.** 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. **TM4 microarray software suite.** Spotfinder. [Online.] http://www.tm4.org /spotfinder.html.

39. **Towbin, H., T. Staehelin, and J. Gordon.** 1992. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. 1979. Biotechnology **24:**145–149.

40. **Tsai, C. M., and C. E. Frasch.** 1982. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. Anal. Biochem. **119:**115–119.

41. **van Beers, S. M., M. Y. de Wit, and P. R. Klatser.** 1996. The epidemiology of Mycobacterium leprae: recent insight. FEMS Microbiol. Lett. **136:**221–230.

42. **World Health Organization.** 2002. World Health Organization Fact sheet 101. Wkly. Epidemiol. Rep. **77:**1–3.

43. **Xiao, Y. D., A. Clauset, R. Harris, E. Bayram, P. Santago II, and J. D. Schmitt.** 2005. Supervised self-organizing maps in drug discovery. 1. Robust behavior with overdetermined data sets. J. Chem. Inf. Model. **45:**1749–1758.