Novel Autoimmune Hepatitis-Specific Autoantigens Identified Using Protein Microarray Technology

ARTICLE *in* JOURNAL OF PROTEOME RESEARCH · JULY 2009

Impact Factor: 4.25 · DOI: 10.1021/pr900131e · Source: PubMed

Published in final edited form as:
J Proteome Res. 2010 January ; 9(1): 30–39. doi:10.1021/pr900131e.

Novel Autoimmune Hepatitis-Specific Autoantigens Identified Using Protein Microarray Technology

Qifeng Song[1,†,§], **Guozhen Liu**[4,5,§], **Shaohui Hu**[1,3], **Yan Zhang**[2], **Yong Tao**[1], **Yuning Han**[1], **Haipan Zeng**[4], **Wei Huang**[1], **Fang Li**[1], **Peng Chen**[4], **Jianhui Zhu**[1], **Chaojun Hu**[2], **Shulan Zhang**[2], **Yongzhe Li**[2,#], **Heng Zhu**[1,3,#], and **Lin Wu**[1,*]

[1]Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 101318, China

[2]Department of Department of Rheumatology, Peking Union Medical College Hospital, Peking Union Medical College & Chinese Academy of Medical Sciences, Beijing 100730, China

[3]Department of Pharmacology & Molecular Sciences and the HiT Center, The Johns Hopkins University School of Medicine, Maryland 21205, USA.

[4]Beijing Protein Innovation Co., Ltd., Beijing 101318, China

[5]College of Life Sciences, Agricultural University of Hebei, Baoding, Hebei, China

Abstract

Autoimmune hepatitis (AIH) is a chronic necroinflammatory disease of the liver with a poorly understood etiology. Detection of non-organ-specific and liver-related autoantibodies using immunoserological approaches has been widely used for diagnosis and prognosis. However, unambiguous and accurate detection of the disease requires the identification and characterization of disease-specific autoantigens. In the present study, we have profiled the autoantigen repertoire of patients with AIH versus those with other liver diseases, identifying and validating three novel and highly specific biomarkers for AIH. In Phase I we fabricated a human protein chip of 5,011 non-redundant proteins and used it to quickly identify 11 candidate autoantigens with relative small serum collection. In Phase II we fabricated an AIH-specific protein chip and obtained autoimmunogenic profiles of serum samples from 44 AIH patients, 50 healthy controls, and 184 additional patients suffering from hepatitis B, hepatitis C, systemic lupus erythematosus, primary Sjögren's syndrome, rheumatoid arthritis, or primary biliary cirrhosis. Using this two-phase approach, we identified three new antigens, RPS20, Alba-like, and dUTPase, as highly AIH-specific biomarkers, with sensitivities of 47.5% (RPS20), 45.5% (Alba-like), and 22.7% (dUTPase). These potential biomarkers were further validated with additional AIH samples in a double-blind design. Finally, we demonstrated that these new biomarkers could be readily applied to ELISA-based assays for use in clinical diagnosis/prognosis.

*Correspondence: Dr. Lin Wu, Beijing Institute of Genomics, Chinese Academy of Sciences, B-6 Airport Industrial Zone, Beijing 101318, China, E-mail: wul@genomics.org.cn, Tel: +86-10-80481195, Fax: +86-10-80485460. #Additional corresponding authors: Dr. Heng Zhu, E-mail: Heng.Zhu@jhmi.edu; Professor Yongzhe Li, E-mail: yongzhelipumch@yahoo.com.cn.
§These authors contributed equally.
†Current address: Leeds Institute of Molecular Medicine, Section of Experimental Therapeutics, Wellcome Trust Brenner Building, St James's University Hospital, Leeds LS9 7TF, UK

SUPPORTING INFORMATION AVAILABLE

Table S1 The list of 5,011 proteins on the microarray is available free of charge via the Internet at http://pubs.acs.org.

Keywords

AIH; biomarker; human protein chip; autoantigens; serum; autoimmune; human liver; clinical proteomics

INTRODUCTION

Autoimmune hepatitis (AIH) is a progressive, necroinflammatory liver disease of cryptic etiology[1] that affects people regardless of age and gender; however, some people are known to be genetically susceptible to this disease. Evidence suggests that the condition is associated with the human complement allele C4AQO and HLA haplotypes B8, B14, DR3, DR4, and Dw32. Nevertheless, the pathogenesis of the condition remains unclear. Patients exposed to effective treatment early in the course of the disease may have their symptoms suspended; thus, early and precise diagnosis is clearly important[3]. Despite the difficulty involved in differentiating AIH from non-AIH cases, diagnostic criteria were codified by an international panel in 1993 and revised in 1999[4, 5]. One of the significant parameters in this scoring system is the presence of autoantibodies in the patient's serum.

AIH has been divided into two types that differ in their immunoserological presentation. The AIH-1 type is characterized by the presence of smooth muscle autoantibodies (anti-SMA) and/or antinuclear autoantibodies (ANA). The targets of the anti-SMA autoantibodies are F-actin and actin, and the probable targets of ANA are the histidines, DNA-binding proteins, and the centrosome complex. The AIH-2 type is characterized by the presence of autoantibodies against liver and kidney microsomal antigens (anti-LKM) and/or liver cytosol antigen 16, 7. It was previously thought that a third AIH type also existed, one that was characterized by the presence of anti-soluble liver antigen/liver pancreas (anti-SLA/LP) autoantibodies. However, it has recently been acknowledged that this variety of AIH should be considered an AIH-1 type.

At present, indirect immunofluorescent (IIF) staining of rodent tissue and HEp-2 cells is the most frequently used method for detecting circulating ANA, SMA, and anti-LKM autoantibodies[1]; the antigens recognized by these autoantibodies are used as diagnostic antigens for profiling autoantibodies in patients' sera. ELISA methods utilizing autoantibodies that target specific autoantigens are being used, or are under development, as a complementary approach to IIF diagnosis, with the goal of improving diagnostic accuracy8. As yet, however, there is no compelling evidence regarding the identity of the ANA targets. By screening a human liver lambda gt11 cDNA expression library9, Manns et al. have identified cytochrome P450db1 as one of the target antigens of anti-LKM autoantibodies. Using inhibition ELISAs and serum samples from controls and patients with various liver diseases to screen cDNA expression libraries for SLA autoantibodies, another group has identified a 50-kDa protein, cytosolic UGA-suppressor tRNA associated protein, as the target of the anti-SLA/LP autoantibody[10]. Asialoglycoprotein receptor (ASGP-R) is a liver-specific protein that is localized to the surface of the liver cell; using a radioimmunoassay, researchers have detected immunoreactivity against this protein in antisera from AIH patients[11]. Detection of autoantibodies against ASGP-R has proved to be of great value for the diagnosis of AIH patients who are seronegative for SMA, ANA, and other classes of autoantibodies12.

The use of conventional methods to screen expression libraries with AIH and control sera in order to identify AIH-specific autoantigens is a time-consuming, tedious process that tends to miss proteins that are in low abundance but robust in serological assays for AIH. Protein microarray technology is a newly developed analytical method that has been shown to have great potential for identifying autoimmune disease biomarkers[13]. In a previous publication, we

reported the construction of a protein chip containing more than 1,000 recombinant human liver proteins and described the use of this chip to identify antigens in a high-throughput approach to monoclonal antibody production[14].

We have now expanded the protein collection on this chip to 5,011 human proteins and used this chip to test the hypothesis that a panel of autoantigens can be used as capture molecules in a microarray to improve both the sensitivity and specificity of AIH diagnosis. To assess this hypothesis, we made use of the expanded chip to screen serum samples from healthy serum controls and individuals with AIH as a means of identifying novel AIH-specific autoantigens. In order to validate the specificity of the candidate autoantigens that we had identified in our screening assay, we developed an AIH-related microarray containing all the candidate autoantigens and probed this microarray with serum samples from patients with a variety of liver diseases, including AIH, hepatitis B (HB), hepatitis C (HC), systemic lupus erythematosus (SLE), primary Sjögren's syndrome (pSS), rheumatoid arthritis (RA), and primary biliary cirrhosis (PBC). These microarray results demonstrated that three previously identified autoantigens (p62, CYP2D6, and ASGP-R [H2 subunit]) and three previously unrecognized autoantigens (RPS20, Alba-like protein, and dUTP diphosphatase) could specifically differentiate AIH patients' sera from other sera. Using two algorithms developed in the current study, we were able to demonstrate that the six autoantigens could be used in combination to clearly diagnose AIH-positive serum samples; we further showed that these autoantigens could be effectively used in protein microarray assays as well as traditional ELISA-based assays.

MATERIALS AND METHODS

Serum Samples

Human serum samples were obtained between 2003 and 2006 from 278 individuals at the Peking Union Medical College Hospital: 44 AIH, 50 PBC, 41 HC, 43 HB, 11 SLE, 11 pSS, and 2 RA patients, as well as 26 with non-autoimmune diseases and 50 healthy subjects. All AIH patients were clinically diagnosed according to the scoring system established by the IAIHG in 1999. All of the AIH serum samples had been clinically characterized as positive for ANA (n=28), SMA (n=21) and/or SLA/LP (n=3) by indirect immunofluorescence assay (Euroimmune, Germany). An additional 41 serum samples from 15 AIH patients and 26 healthy controls were collected at Beijing Union Medical College Hospital in 2007 for validation experiments. The study was approved by the IRBs at Peking Union Medical College Hospital.

Human Gene Cloning

Human gene cloning was carried out as previously described14, with one modification: The samples used for the PCR templates were individual cDNA clones that had been sequence-identified and isolated from various human tissues. The resulting PCR products were subcloned into pEGH using a yeast homologous recombination strategy[15]. In brief, primers containing yeast recombinant sites were designed and used to PCR-amplify human coding regions, and the PCR products were co-transfected into yeast (Y258) with the linearized yeast expression vector pEGH containing the homology recombination sites. The resulting recombined plasmids were amplified and carried in the yeast. To monitor the success of the open reading frame (ORF) cloning, the resulting plasmid DNAs were extracted, and the inserts were amplified using generic primers flanking the recombination sites. Insert sizes were determined by gel electrophoresis.

Protein Purification and Fabrication of the Human Protein Chips

Protein expression and purification were conducted as described previously[15]. The purified GST fusion proteins were eluted into printing buffer containing 30% glycerol in 50 mM HEPES (pH 7.0). The eluates were arrayed in 384-well titer dishes for printing. The quality of the

protein purifications was verified by SDS-PAGE. Purified human proteins and control proteins were spotted in duplicate onto poly-L-lysine-coated slides at high density using a SpotArray-72 printer (Perkin Elmer). The printed human protein chips were kept horizontal at room temperature (RT) for 1 h before storage at 4° C.

Serum Assays on Human Protein Chips

Protein chips were first blocked by immersion in blocking buffer (PBST, 0.1% [v/v] PBS with 0.05% [v/v] Tween 20, and 1% [w/v] BSA) at RT for 1 h with gentle shaking. Subsequently, 200 μL of serum (diluted 1000-fold in PBS buffer) was added to the chip and incubated under a glass coverslip (LifterSlip, Erie Scientific Company, Portsmouth, NH, USA) to maintain the appropriate humidity. After an 1-h incubation at RT, the chip was subjected to three 10-min washes with PBST at 40° C, followed by a 1-h incubation with 1,000-fold-diluted anti-human IgG antibody labeled with Cy-5 (The Jackson Laboratory, Bar Harbor, ME, USA) in the dark at RT. The chip was again washed three times with PBST. After two rinses in double-deionized water, the chip was blown to dryness with compressed air and scanned with a BioCapital microarray scanner (BioCapital, Beijing, China). The binding signals were acquired and analyzed using the GenePix Pro 5.0 software (Molecular Devices, Sunnyvale, CA, USA).

Analysis of the Chip Data

Following image analysis with GenePix Pro 5.0, the signal intensity of each protein feature was determined by subtracting the median background value from the median foreground value. When the resulting signals were <0, the protein spots were assigned a minimum signal of 1. Since every protein was spotted in duplicate, the average of each duplicate was used as the final signal intensity for a given protein.

The total signal intensity was used to perform inter-chip normalization. The total signal intensity for a given chip (chip $signal_n$) was divided by the mean value for all the chips (average (chip $signal_n$)) to obtain an intensity factor for the chip (intensity factor ($chip_n$)). Each signal for a given chip was then divided by the intensity factor ($chip_n$) to generate a normalized signal.

The mean value of a given protein for all control serum samples was then calculated and termed the $average_{protein\ i}$ control value. Dividing the signal intensity value for each protein in the individual patient's serum by the $average_{protein\ i}$ control value yielded the so-called "factor F" value. A protein was considered positive when its F value was $>2^{16}$. The number of each protein's positive values for the patient sera was counted, and those proteins with a positive rate >60% for all the patient sera were considered AIH-specific autoantigen candidates.

Fabrication of an AIH-related Protein Microarray and Its Probing with Human Sera

All the AIH-specific autoantigen candidates identified on the human protein chip were printed together with controls (human IgG, glutathione-S-transferase [GST], and printing buffer) in duplicate within a 2.52 mm×1.44 mm area, and 12 identical probe areas were fabricated on a single poly-L-lysine-treated microscope slide. The printed microarrays were allowed to remain at RT for 1 h before storage at 4 ° C.

In order to probe the protein microarrays with human sera, a 12-hole rubber gasket (BioCapital, Beijing, China) was applied to each microarray to form 12 individual chambers. The probing procedure was basically identical to that performed without chambers, except that 50 μl of 1/1,000-diluted human serum was incubated in each microarray chamber. After the incubation, the microarrays were rinsed with wash buffer before the rubber gaskets were carefully removed. The chips were scanned using a BioCapital microarray scanner (BioCapital), and the signals were analyzed using GenePix Pro 5.0 software (Molecular Devices).

Microarray Data Analyses to Identify AIH-Specific Autoantigens

The mean signal intensities and standard deviations (SD) for 50 healthy controls were calculated, and the value for the $\text{mean}_{\text{healthy}}+3\times\text{SD}$ for each particular potential antoantigen was used as an immnoreactive cutoff value.

To determine the specificity and resolution of each of the potential AIH autoantigens, we applied two commonly used computational methods, logistic regression and discriminant analysis[17], using open-source SPSS 12.0 software. In the case of both methods, two group analyses were conducted, for the AIH and non-AIH groups. Logistic regression is a generalized linear regression model for binary responses. The candidate features were selected by the model using a dual-direction stepwise search with Akaike's information criteria, and the probabilities were calculated as follows: $p = \text{EXP}(\Sigma(b_i x_i)+c)/(1+\text{EXP}(\Sigma(b_i x_i)+c))$, where p is the probability of each case, i=1 to n; b is the regression coefficient of a given autoantigen; × is the signal intensity; and c is a constant generated by the model. For prediction, only those cases with a >0.5 probability were classified as AIH-positive. The model first performed t-tests for each of the 14 autoantigens and selected those with p values <0.01 for the model building.

Discriminant analysis is a linear regression model that predicts group membership by using a set of predictors. The formula is $y=\Sigma(b_i x_i)+c$, where i=1 to n, x is the signal intensity, b represents the unstandardized canonical discriminant function coefficient, and c is a constant provided by the model. Those cases in which the y value was >0 were considered AIH-positive. The model first generated standardized canonical discriminant function coefficients for each autoantigen in order to evaluate its importance in prediction. The selected autoantigens were then used for model building.

ELISA Assays

Purified recombinant protein (100 µl at 2.5 µg/ml) was added to the wells of a 96-well plate and incubated for 1 h at 37 °C. The wells were blocked by discarding the protein solution and adding 200µl of blocking buffer (1% BSA), then incubating the samples for 1 h at 37 °C. The blocking buffer was then discarded, and the sera (1:40 dilutions) were individually added to the wells and incubated for 1 h at 37 °C. The serum dilutions were removed, and the wells were washed three times with 300 µl PBST each time; HRP-conjugated mouse anti-human IgG antibody (diluted 1:5,000) was then added, and the samples were incubated for 1 h at 37°C. The wells were again washed three times with 300 µl PBST each before the addition of 100 µl tetramethylbenzidine (TMB) substrate solution (10 µg TMB, 1 µl DMSO, 0.51 g trisodium citrate dihydrate, 0.184 g disodium hydrogen phosphate dodecahydrate, 10 µl H_2O_2, and 99 µl H_2O). After a 5- to 10-min incubation at RT, the reaction was stopped by adding 100 µl sulfuric acid (4.5 N) per well, and the absorption at 450 nm (OD_{450}) was measured using an automatic ELISA reader (Tianshi, Beijing, China).

RESULTS

Construction of a Protein Chip of 5,011 Unique Human Proteins

We have previously reported the fabrication of a human protein microarray containing 1,058 liver-expressing proteins in full-length and demonstrated its application in high-throughput antigen identification[14]. To enable a broader use of this technology platform, particularly with regard to autoantigen identification, we first attempted to increase the protein content of the chips by adding proteins expressed in a wide range of tissues. First, we designed a primer set for PCR amplification of 5,670 human ORFs using cDNA templates obtained from human fetal brain, bone marrow, pancreas, and testicle as well as adult and fetal liver. Second, we used a yeast recombinant cloning strategy to clone these ORFs into a yeast expression vector, pEGH. A total of 4,053 ORFs were successfully cloned. Together with the 1,058 previously

reported proteins, a total of 5,011 non-redundant human proteins (88.6 % full-length) were over-expressed and purified as GST fusion proteins in yeast.

In a pilot assay, we found that poly-L-lysine-coat slides produced a signal-to-background ratio comparable to that of other surfaces, such as FAST and Fullmoon, in serum profiling assays (data not shown). At the same time, we also tested a variety of different conditions in order to optimize the serum profiling assay, including a range of serum concentrations (200-to 5000-fold dilutions) and washing conditions (e.g., high salt versus low salt, and room temperature versus 40°C). On the basis of these experiments, we chose poly-L-lysine slides for chip fabrication, 1000-fold dilution as the most appropriate serum concentration, and 40°C as the chip-washing temperature (see the Materials and Methods for details).

We then fabricated a second generation of human protein chips by printing all the purified proteins onto poly-L-lysine-coated microscope slides. Because CYP2D6 and ASGP-R had already been identified as autoantigens in AIH patients[18, 19], and p62 has been implicated in some PBC cases[20], we included these three proteins on the chips as positive controls. Human IgG at a known concentration was also printed at the upper left corner of each block on the chips to serve as a control and landmark. The quality and quantity of the immobilized proteins on the chips were determined by probing with an anti-GST antibody, and 84.5% of the proteins produced signals that were significantly above background (Fig. 1).

Autoantibody Profiling to Identify AIH-Specific Autoantigens

To save time and expense with regard to the serum profiling assays, we carried out the assays in two phases: In Phase I, we selected 22 AIH and 30 control serum samples and probed the human protein chips with the samples individually at a 1000-fold dilution, followed by detection of bound human autoantibodies using a Cy-5-conjugated anti-human IgG antibody. We observed that a large number of proteins (e.g., in the hundreds) could be readily recognized by sera from both patients and healthy controls, and the serum profiles also showed obvious individual-to-individual variation within both the patient and control groups (Fig. 2), a phenomenon that has been reported in previous investigations[21, 22].

To identify potential autoantigens that are specifically associated with AIH, we used GenePix software to acquire and process the protein chip data. Binding signals for each serum probing were first normalized using the overall chip signals (as detailed in Materials and Methods). Then, the value of each signal produced by a protein that reacted with the AIH serum was divided by the mean value of the protein's signal from the healthy control serum samples (negative controls). The resulting ratio was defined as factor F, reflecting the differential autoimmunogenic potential of a given human protein, or autoantigen. At F values >2, all three positive control antigens (p62, CYP2D6, and ASGP-R) identified at least 13 of the 22 AIH-positive serum samples. Therefore, we decided to use this criterion to score the potential autoantigens. This scoring system identified 11 candidate autoantigens (Table 1), of which two (RPS20 and Alba-like) were ranked higher than all of the three previously known antigens. None of the patients showed a positive reaction with all of the candidate antigens.

In Phase II, we fabricated an AIH-specific protein chip that carried the 11 candidate autoantigens and the three positive control antigens, in addition to GST and BSA as negative controls. This strategy allowed us to fabricate protein chips in large number at low cost and to screen a large number of serum samples in a relatively short time (Fig. 3). To determine which candidate autoantigens were AIH-specific, we sequentially probed the AIH-specific chips with the serum samples used in Phase I and obtained data for samples from an additional 22 AIH, 50 PBC, 43 HB, 41 HC, 11 SLE, 11 pSS, and 2 RA patients using the same protocol as described above. As negative controls, we also included 26 serum samples from patients suffering from

other types of severe diseases and 50 samples from healthy subjects. Representative probing images are shown in Fig. 3.

Identification of Novel AIH-specific Autoantigens

After the same process of data normalization, we identified AIH-specific autoantigens by using a calculation that was more stringent than the one used in Phase I. We first calculated the mean value ($mean_{healthy}$) and standard deviation (SD) for the individual signal intensities obtained for the 50 healthy control sera. Using a value of three times the SD above $mean_{healthy}$ as a cutoff value, we scored the positives for each serum probing reaction and calculated the true- and false-positive values for each antigen presented on the chips (Table 2). Three known antigens (p62, CY2D6, and ASGP-R) and three new candidate antigens (RPS20, Alba-like, and dUTP diphosphatase [dUTPase]), showed a true positive value for the AIH samples that ranged from 20.5% to 47.7%, significantly higher than the values for the remaining eight candidate autoantigens. These six antoantigens also showed a slight cross-reactivity with PBC, HB, HC, and other control samples (Table 2), which is consistent with the behavior of autoantigens reported previously[23, 24].

We next applied a logistic regression algorithm, a generalized linear regression analysis for binary responses, to analyze the serum profiling data. Of the 11 candidate and 3 known autoantigens, the top 6 selected by the software because of their low *p* values (<0.01) in t-tests were the same 6 antigens identified in the screening described above (Table 2). Using the signals for these six proteins, we created a logistic regression model. The formulas for the logistic regression and *p* value calculation were (Formula I):

$$y=0.0054X_1+0.0035X_2-0.00052X_3+0.000186X_4-0.0083X_5+0.000284X_6-5.2$$

$$p=\mathrm{EXP}(y)/(1+\mathrm{EXP}(y)),$$

where X_1 to X_6 represent the signals from RPS20, Alba-like, p62, CYP2D6, dUTPase, and ASGP-R, respectively. When our model was used to classify the 278 training samples, it achieved an accuracy of 91.7% (sensitivity of 68.2% and specificity of 96.2%, with the sensitivity and specificity defined as correctly classified positive and negative sera, respectively) (Table 3).

We next calculated the standardized canonical discriminant function coefficients and found that seven autoantigens had a positive contribution, and the six antigens identified by the microarray analysis and logistic regression algorithm were among them (Table 2). Using the signals for these six proteins, we created a second discriminant model, Formula II:

$$y=0.0033X_1+0.0013X_2+0.0001X_3-0.0005X_4-0.003X_5+2.73292E{-}05X_6-1.41,$$

where X_1 to X_6 represent the signals from RPS20, Alba-like, p62, CYP2D6, dUTPase, and ASGP-R, respectively. When the *y* value was >0, the autoantigen was scored as positive. Using this model, we were able to classify the AIH-positive training cases versus the non-AIH cases with an overall accuracy of 91.7% (sensitivity of 84.1% and specificity of 93.2%) (Table 3).

These analyses indicated that both computational models identified RPS20, Alba-like, p62, CYP2D6, ASGP-R (H2 subunit), and dUTP diphosphatase as powerful potential autoantigens for use in differentiating AIH from non-AIH human serum samples, consistent with the results

we had obtained by calculating positives using cutoff value. Thus, our combined analyses convincingly demonstrate the specificity of these six proteins for AIH.

Finally, we conducted box-whisker plot analysis to visualize the range of signal intensities for the six candidate autoantigens across the various groups of sera. Two autoantigens (RPS20 and Alba-like) identified in this study were not only able to differentiate AIH from non-AIH sera but also to differentiate AIH from PBC or HCV serum samples (Fig. 4). Because the previously identified autoantigens used for AIH diagnosis are often detected in sera from PBC and HCV patients[23], RPS20 and Alba-like offer the important advantage that they can be used as more specific autoantigen biomarkers for diagnosing AIH.

Validation of the Protein Microarray Approach for AIH Diagnosis

In order to validate the specificity of the three autoantigens identified using our protein microarray approach and to assess their utility for AIH diagnosis, we conducted a double-blind experiment using 41 additional serum samples that were collected at the Beijing Union Medical College Hospital in 2007; 15 of these samples had been identified as AIH-positive. Using the protocol described above, we obtained the autoantigen profiles of all 41 serum samples using the AIH-specific chips. When we applied the logistic regression model, 9 of the15 AIH sera were diagnosed as positive, whereas none of the healthy controls were scored as AIH-positive. With the discriminant model, we were able to recognize 10 of the 15 samples as AIH sera and, again, none of the healthy controls were misdiagnosed. Therefore, the protein microarray approach was confirmed as being able to diagnose AIH with an accuracy of 85.4% (logistic regression model: 60% sensitivity and 100% specificity) or 87.8% (discriminant model: 66.7% sensitivity and 100% specificity) (Table 4).

Conversion to Diagnostic Biomarkers

To determine whether the autoantigens identified by the protein chip approach could be readily transformed for use as diagnostic biomarkers for AIH in a traditional ELISA format, we tested the three new autoantigens (*i.e.*, RPS20, Alba-like, and dUTPase) in an ELISA assay and compared the results we obtained to those obtained for the three known autoantigens (CYP2D6, ASGPR2, and p62), using 15 AIH-positive and 15 healthy serum samples.

We then calculated the mean value and standard deviation of the OD_{450} values for the healthy serum samples in the ELISA tests. When we used a value of three times the SD of the negative controls as a cutoff, we found that 9 of the 15 AIH-positive sera were positive for at least one of the three new autoantigens, and 10 of 15 reacted with at least one of the three previously known autoantigens (Fig. 5). Four AIH-positive samples did not react with either of these groups of autoantigens. Consistent with previous reports, p62 had the lowest sensitivity (2 of 15); however, sample AIH10 was only reactive with p62, indicating that it is still useful as a biomarker. The other five autoantigens showed comparable sensitivities, ranging from 26.7% to 46.7%. It is also noteworthy that sample AIH7 was reactive only with a new biomarker, dUTPase. In summary, the new autoantigens can be readily applied to a traditional ELISA-based clinical diagnostic format and, together with the existing biomarkers, can be used to improve the accuracy and sensitivity of AIH diagnosis.

DISCUSSION

Proteomics technologies have been employed to profile autoimmune disease samples in order to identify specific autoantibodies for diagnosis and/or prognosis[25]. In recent years, we have witnessed a rapid development of the protein chip approach in this fast-moving field. Robinson and colleagues were among the first to develop autoanigen microarrays containing hundreds of known autoantigens that could be used to profile autoimmune diseases (not including AIH)

[26]. Others have screened random cDNA or phage-display libraries in a protein chip format[27–30]. Recently, Snyder and colleagues identified four biomarkers for ovarian cancer using a commercially available human protein chip of 5,005 proteins[31]. However, these approaches, although fruitful, have proved to be time-consuming, insufficiently comprehensive, or high in cost.

We now report the development of a new strategy that has allowed us to quickly narrow down the number of potential autoantigens to a manageable number. Our two-phase strategy involved probing a high-content human protein chip with fewer than two dozen patient serum samples, and then following this preliminary step with extensive serum profiling using a disease-specific (AIH) protein chip. By taking advantage of existing biomarkers, we were able to readily identify 11 potential autoantigens in Phase I, and then further confirm three of these proteins as newly identified AIH-specific autoantigens in Phase II in experiments using a large number of AIH-positive and control sera. Using two kinds of statistical algorithms, we were able to achieve similar diagnostic results, with high sensitivity and specificity. The effectiveness of our strategy was then validated using additional AIH samples in a double-blind format. Thus, our analyses confirm that this new strategy has considerable potential as an efficient and cost-effective method for identifying relevant biomarkers of AIH and, by extension, of other diseases.

The new autoantigens identified in this study may provide additional clues regarding the composition of anti-nuclear autoantibodies (ANAs). The Alba-like protein is a member of the Alba family, which includes the archaeal protein Alba and a number of eukaryotic proteins with no known function. The DNA/RNA-binding protein Alba binds double-stranded DNA tightly but without sequence specificity. It binds rRNA and mRNA *in vivo* and may play a role in maintaining the structural and functional stability of RNA and, perhaps, ribosomes. It is distributed uniformly and abundantly throughout the chromosomes32. The dUTP diphosphatase (dUTPase) has been identified in the form of two splice variants, with dUTP-N localized to the nuclei in HeLa cells33. Hence, both proteins are presumably representative autoantigen targets among the ANAs. Interestingly, the ribosomal protein RPS20 has previously been identified as an autoantigen in SLE patients, but with a relatively low sensitivity of detection (20%)34. In the present study, we also observed that one out of 11 serum samples (9%) from SLE patients recognized RPS20; however, this protein was a much more sensitive autoantigen (48%) in the case of the AIH patients in our study (Table 2).

CONCLUSIONS

Our new approach and relevant clinical findings provide strong evidence that advances in high-throughput protein microarray technology have the potential to rapidly identify novel biomarkers useful for autoimmune diseases and to improve the diagnosis and/or prognosis of AIH and other types of autoimmune disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by funding from Ministry of Science and Technology of the People's Republic of China (2006AA02A311 to L.W.), National Natural Science Foundation of China (30640084 and 30872331 to Y.L.), and the NIH (R01GM076102 to H.Z.). We thank Dr. Deborah McClellan for editorial assistance.

REFERENCES

1. Zachou K, Rigopoulou E, Dalekos GN. Autoantibodies and autoantigens in autoimmune hepatitis: important tools in clinical practice and to study pathogenesis of the disease. J Autoimmune Dis 2004;1:2. [PubMed: 15679907]
2. Czaja AJ, Donaldson PT. Genetic susceptibilities for immune expression and liver cell injury in autoimmune hepatitis. Immunol Rev 2000;174:250–259. [PubMed: 10807521]
3. Schvarcz R, Glaumann H, Weiland O. Survival and histological resolution of fibrosis in patients with autoimmune chronic active hepatitis. J Hepatol 1993;18:15–23. [PubMed: 7688011]
4. Johnson PJ, McFarlane IG. Meeting report: International Autoimmune Hepatitis Group. Hepatology 1993;18:998–1005. [PubMed: 8406375]
5. Alvarez F, Berg PA, Bianchi FB, Bianchi L, Burroughs AK, Cancado EL, Chapman RW, Cooksley WG, Czaja AJ, Desmet VJ, Donaldson PT, Eddleston AL, Fainboim L, Heathcote J, Homberg JC, Hoofnagle JH, Kakumu S, Krawitt EL, Mackay IR, MacSween RN, Maddrey WC, Manns MP, McFarlane IG, Meyer zum Büschenfelde KH, Mieli-Vergani G, Nakanuma Y, Nishioka M, Penner E, Porta G, Portmann BC, Reed WD, Rodes J; Schalm SW, Scheuer PJ, Schrumpf E, Seki T, Toda G, Tsuji T, Tygstrup N, Vergani D, Zeniya M. International Autoimmune Hepatitis Group Report: review of criteria for diagnosis of autoimmune hepatitis. J Hepatol 1999;31:929–938. [PubMed: 10580593]
6. Bridoux-Henno L, Maggiore G, Johanet C, Fabre M, Vajro P, Dommergues JP, Reinert P, Bernardxz O. Features and outcome of autoimmune hepatitis type 2 presenting with isolated positivity for anti-liver cytosol antibody. Clin Gastroenterol Hepatol 2004;2:825–830. [PubMed: 15354284]
7. Duchini A, McHutchison JG, Pockros PJ. LKM-positive autoimmune hepatitis in the western United States: a case series. Am J Gastroenterol 2000;95:3238–3241. [PubMed: 11095348]
8. Bogdanos DP, Invernizzi P, Mackay IR, Vergani D. Autoimmune liver serology: Current diagnostic and clinical challenges. World J Gastroenterol 2008;14:3374–3387. [PubMed: 18528935]
9. Manns MP, Johnson EF, Griffin KJ, Tan EM, Sullivan KF. Major antigen of liver kidney microsomal autoantibodies in idiopathic autoimmune hepatitis is cytochrome P450db1. J Clin Invest 1989;83:1066–1072. [PubMed: 2466049]
10. Wies I, Brunner S, Henninger J, Herkel J, Kanzler S, Meyer Zum Meyer zum Büschenfelde KH, Lohse AW. Identification of target antigen for SLA/LP autoantibodies in autoimmune hepatitis. Lancet 2000;355:1510–1515. [PubMed: 10801173]
11. McFarlane BM, McSorley CG, McFarlan IG, Williams Rl. A radioimmunoassay for detection of circulating antibodies reacting with the hepatic asialoglycoprotein receptor protein. J Immunol Methods 1985;77:219–228. [PubMed: 3981001]
12. Treichel U, Poralla T, Hess G, Manns M, Meyer Zum Büschenfelde KH. Autoantibodies to human asialoglycoprotein receptor in autoimmune-type chronic hepatitis. Hepatology 1990;11:606–612. [PubMed: 1691732]
13. Kalbas M, Lueking A, Kowald A, Muellner S. New analytical tools for studying autoimmune diseases. Curr Pharm Des 2006;12:3735–3742. [PubMed: 17073672]
14. Hu S, Li Y, Liu G, Song Q, Wang L, Han Y, Zhang Y, Song Y, Yao X, Tao Y, Zeng H, Wang J, Zhu H, Chen Z-N, Wu L. A protein chip approach for high-throughput antigen identification and characterization. Proteomics 2007;7:2151–2161. [PubMed: 17549792]
15. Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, Bertone P, Lan N, Jansen R, Bidlingmaier S, Houfek T, Mitchell T, Miller P, Dean RA, Gerstein M, Snyder M. Global analysis of protein activities using proteome chips. Science 2001;293:2101–2105. [PubMed: 11474067]
16. Lueking A, Luekinga A, Huberb O, Wirthsd C, Schultea K, Stielerf KM, Blume-Peytavif U, Kowaldg A, Hensel-Wiegeld K, Tauberd R, Lehrachg H, Meyera HE, Cahilla DJ. Profiling of alopecia areata autoantigens based on protein microarray technology. Mol Cell Proteomics 2005;4:1382–1390. [PubMed: 15939964]
17. Press SJ, Wilson S. Choosing between logistic regression and discriminant analysis. Journal of the American Statistical Association 1978;73:699–705.
18. Yamamoto AM, Mura C, Morales MG, Bernard O, Krishnamoorthy R, Alvarez F. Study of CYP2D6 gene in children with autoimmune hepatitis and P450 IID6 autoantibodies. Clin Exp Immunol 1992;87:251–255. [PubMed: 1346512]

19. O'Brien CJ, Vento S, Donaldson PT, McSorley CG, McFarlane IG, Williams R, Eddleston AL. Cell-mediated immunity and suppressor-T-cell defects to liver-derived antigens in families of patients with autoimmune chronic active hepatitis. Lancet 1986;1:350–353. [PubMed: 2868297]
20. Nesher G, Margalit R, Ashkenazi YJ. Anti-nuclear envelope antibodies: Clinical associations. Semin Arthritis Rheum 2001;30:313–320. [PubMed: 11303304]
21. Nobrega A, Haury M, Grandien A, Malanchere E, Sundblad A, Coutinho A. Global analysis of antibody repertoires. II. Evidence for specificity, self-selection and the immunological "homunculus" of antibodies in normal serum. Eur J Immunol 1993;23:2851–2859. [PubMed: 8223861]
22. Cohen IR. The cognitive paradigm and the immunological homunculus. Immunol Today 1992;13:490–494. [PubMed: 1463581]
23. Miyakawa H, Kikazawa E, Abe K, Kikuchi K, Fujikawa H, Matsushita M, Kawaguchi N, Morizane T, Ohya K, Kako M. Detection of anti-LKM-1(anti-CYP2D6) by an enzyme-linked immunosorbent assay in adult patients with chronic liver diseases. Autoimmunity 1999;30:107–114. [PubMed: 10435724]
24. Michalska Z, Stalke P, Witczak-Malinowska K, Lakomy EA, Sikorska K, Radowska D, Stolarczyk J, Trocha H. Autoimmune reactions in HBV and HCV. Med Sci Monit 2001;7:175–180. [PubMed: 12211715]
25. Hueber W, Utz PJ, Steinman L, Robinson WH. Autoantibody profiling for the study and treatment of autoimmune disease. Arthritis Res 2002;4:290–295. [PubMed: 12223102]
26. Robinson WH, DiGennaro C, Hueber W, Haab BB, Kamachi M, Dean EJ, Fournel S, Fong D, Genovese MC, Neuman de Vegvar HE, Skriner K, Hirschberg DL, Morris RI, Muller S, Pruijn GJ, van Venrooij WJ, Smolen JS, Brown PO, Steinman L, Utz PJ. Autoantigen microarrays for multiplex characterization of autoantibody responses. Nat Med 2002;8:295–301. [PubMed: 11875502]
27. Wang X, Yu J, Sreekumar A, Varambally S, Shen R, Giacherio D, Mehra R, Montie JE, Pienta KJ, Sanda MG, Kantoff PW, Rubin MA, Wei JT, Ghosh D, Chinnaiyan AM. Autoantibody signatures in prostate cancer. N Engl J Med 2005;353:1224–1235. [PubMed: 16177248]
28. Lueking A, Possling A, Huber O, Beveridge A, Horn M, Eickhoff H, Schuchardt J, Lehrach H, Cahill DJ. A nonredundant human protein chip for antibody screening and serum profiling. Mol Cell Proteomics 2003;2:1342–1349. [PubMed: 14517340]
29. Gutjahr C, Murphy D, Lueking A, Koenig A, Janitz M, O'Brien J, Korn B, Horn S, Lehrach H, Cahill DJ. Mouse protein arrays from a TH1 cell cDNA library for antibody screening and serum profiling. Genomics 2005;85:285–296. [PubMed: 15718096]
30. Horn S, Lueking A, Murphy D, Staudt A, Gutjahr C, Schultek K, König A, Landsberger M, Lehrach H, Felix SB, Cahill DJ. Profiling humoral autoimmune repertoire of dilated cardiomyopathy (DCM) patients and development of a disease-associated protein chip. Proteomics 2006;6:605–613. [PubMed: 16419013]
31. Hudson ME, Pozdnyakova I, Haines K, Mor G, Snyder M. Identification of differentially expressed proteins in ovarian cancer using high-density protein microarrays. Proc Natl Acad Sci U S A 2007;104:17494–17499. [PubMed: 17954908]
32. Aravind L, Iyer LM, Anantharaman V. The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism. Genome Biol 2003;4:R64. [PubMed: 14519199]
33. Ladner RD, McNulty DE, Carr SA, Roberts GD, Caradonna SJ. Characterization of distinct nuclear and mitochondrial forms of human deoxyuridine triphosphate nucleotidohydrolase. J Biol Chem 1996;271:7745–7751. [PubMed: 8631816]
34. Kemp EH, Herd LM, Watermana EA, Wilsonb AG, Weetmana AP, Watsona PF. Immunoscreening of phage-displayed cDNA-encoded polypeptides identifies B cell targets in autoimmune disease. Biochem Biophys Res Commun 2002;298:169–177. [PubMed: 12379236]

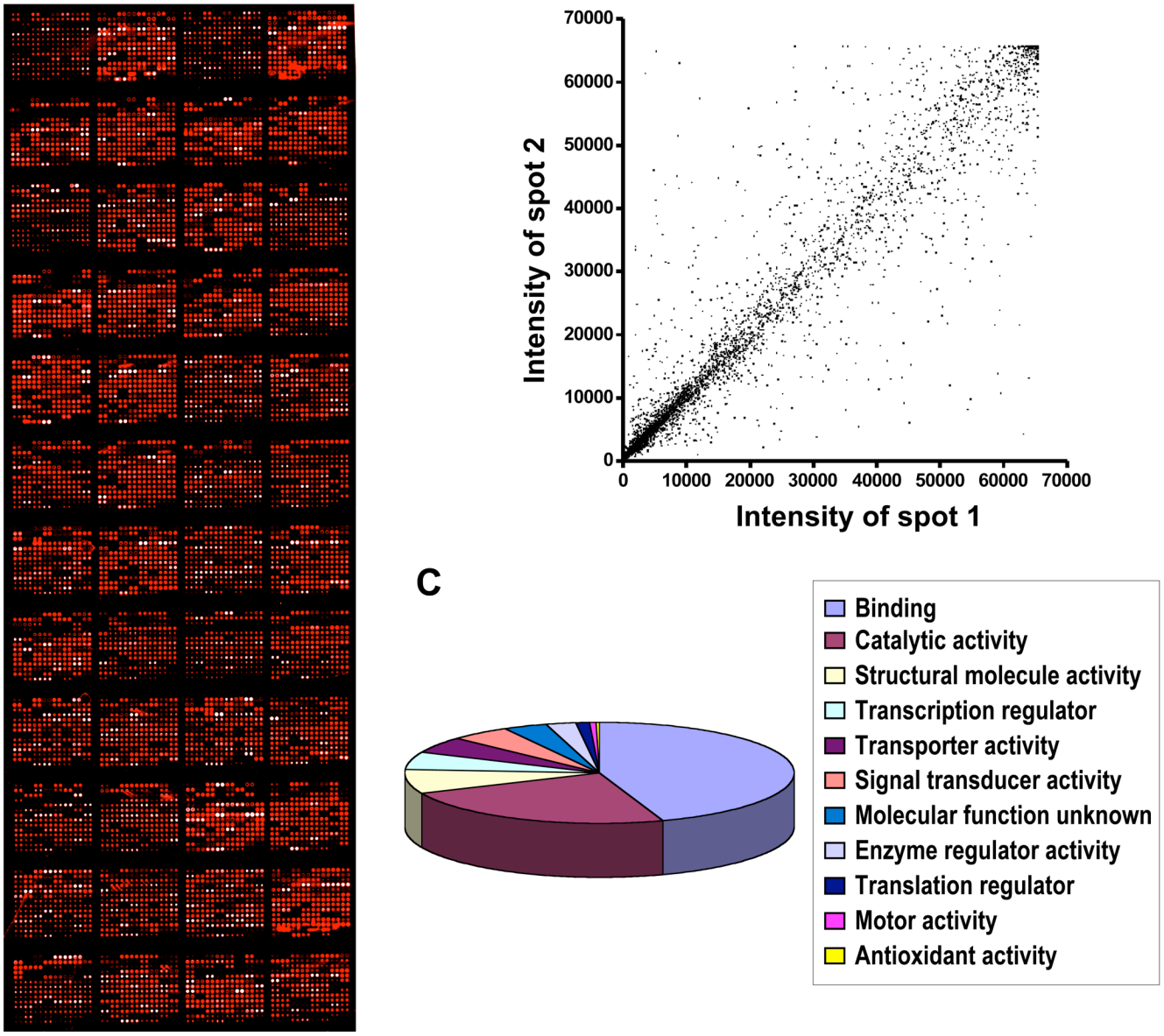

Figure 1. Construction of a human protein chip of 5,011 proteins

(A) The protein chip probed with anti-GST monoclonal antibody. Recombinant human proteins were purified and printed in duplicate on poly-L-lysine coated microscope slides. To monitor the quality and relative quantity of the printed proteins on glass slides, the human protein chips were probed with anti-GST antibody, followed by Cy5-labeled secondary antibody to visualize the signals. The proteins positively detected by the anti-GST antibody are represented in red or white (saturated intensity). (B) Correlation of spot intensities of all the duplicate pairs. The signal intensities of duplicate spots (Spot 1 versus its corresponding Spot 2) were plotted against each other. The resulting correlation coefficient was 0.95, indicating high reproducibility of the protein spotting. (C) Functional distribution of the printed human proteins. The functions were annotated using the gene ontology method.

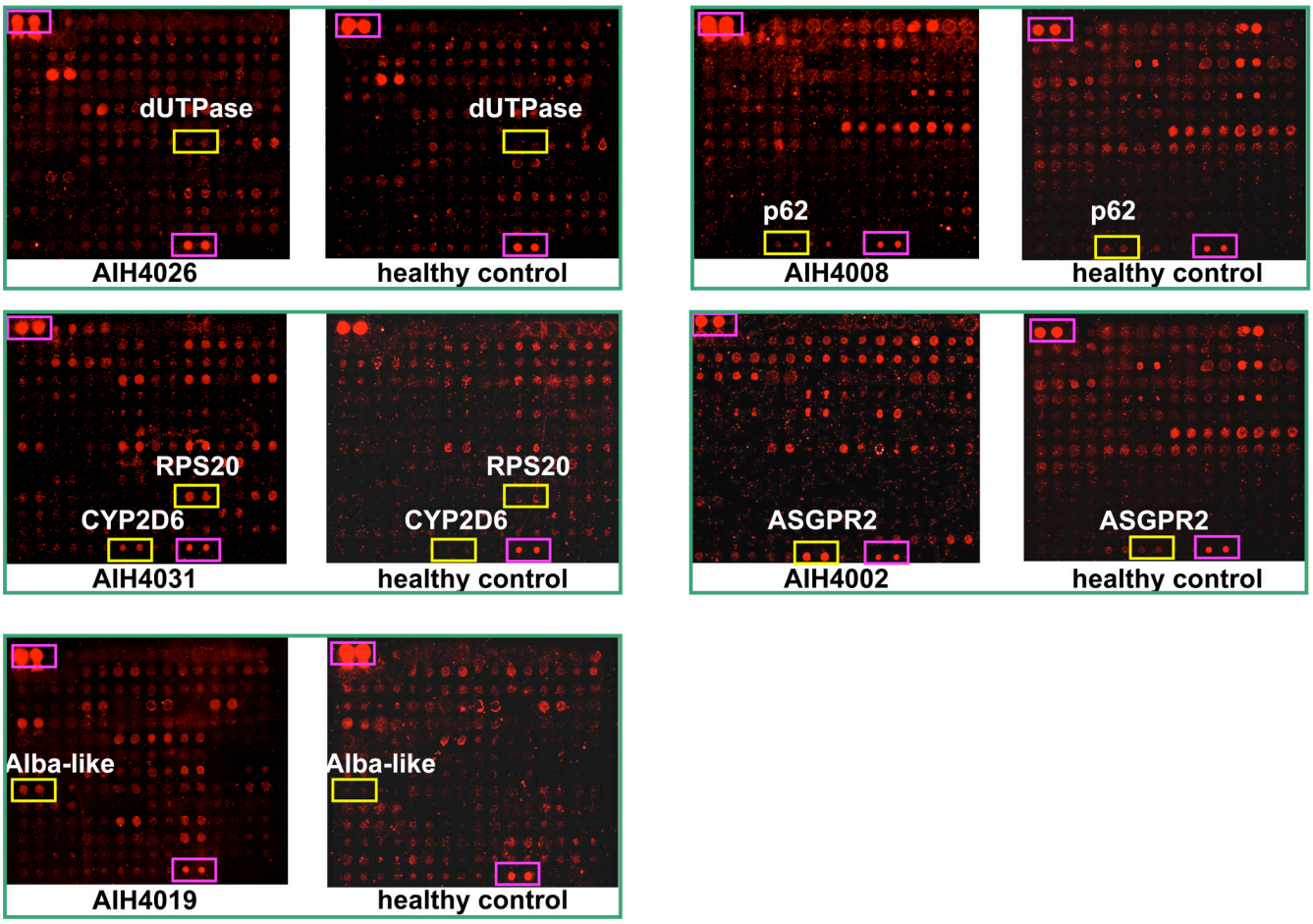

Figure 2. Probing of the human protein chip with AIH-positive and healthy control sera

Twenty-two AIH and 30 healthy control serum samples were diluted 1:1000 and individually incubated with the human protein chip, followed by the addition of the anti-human IgG antibody (Cy5-conjugated). Chips were dried and scanned to acquire the images. Representative areas of the images are illustrated. Yellow boxes indicate the positive candidate autoantigens, and purple boxes indicate human IgG (the positive control).

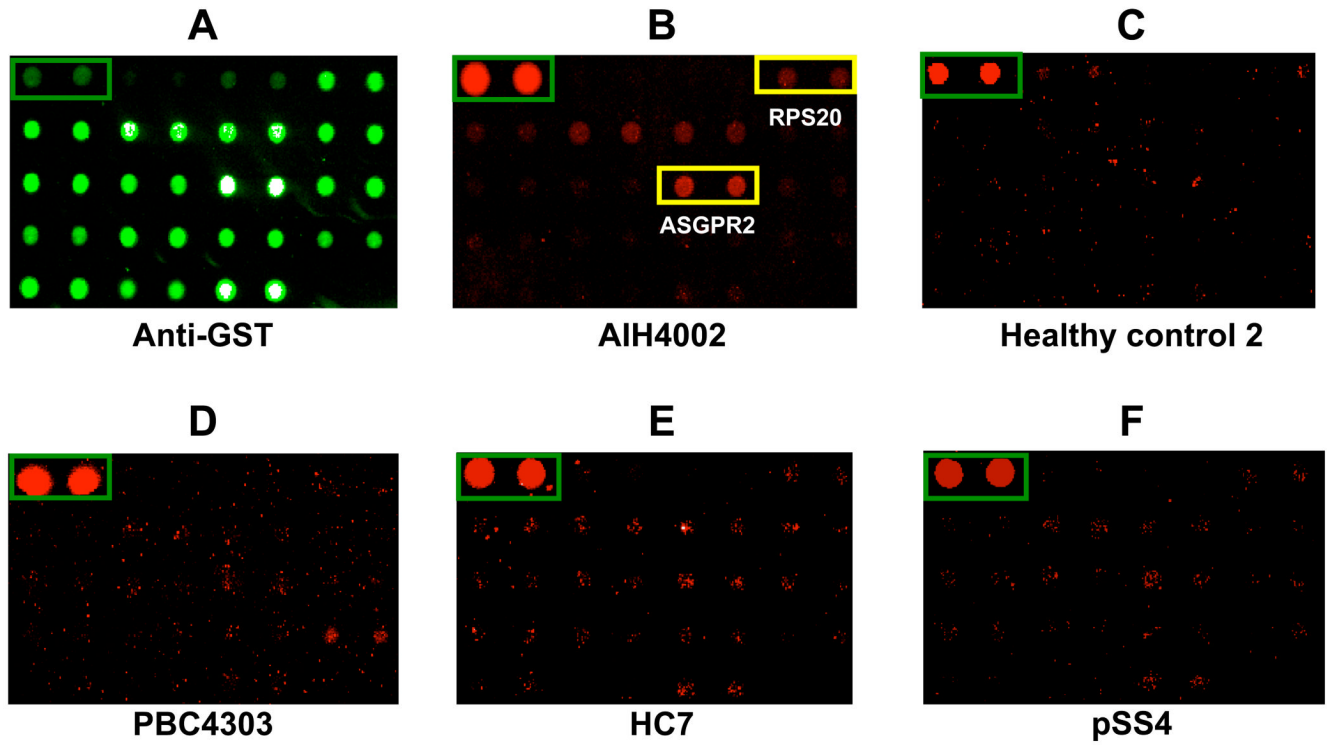

Figure 3. Large-scale serum profiling using AIH-specific chips

Fourteen autoantigen candidates were purified and printed onto a poly-L-lysine coated microscope slide to form the AIH-specific protein chip. Shown are representative chip images probed with either anti-GST antibodies (A) or with sera collected from AIH patients (B), healthy subjects (C), PBC (D), HC (E), or pSS (F) patients. Green boxes show the human IgG spots as a landmark. The yellow boxes indicate the AIH-specific positive features.

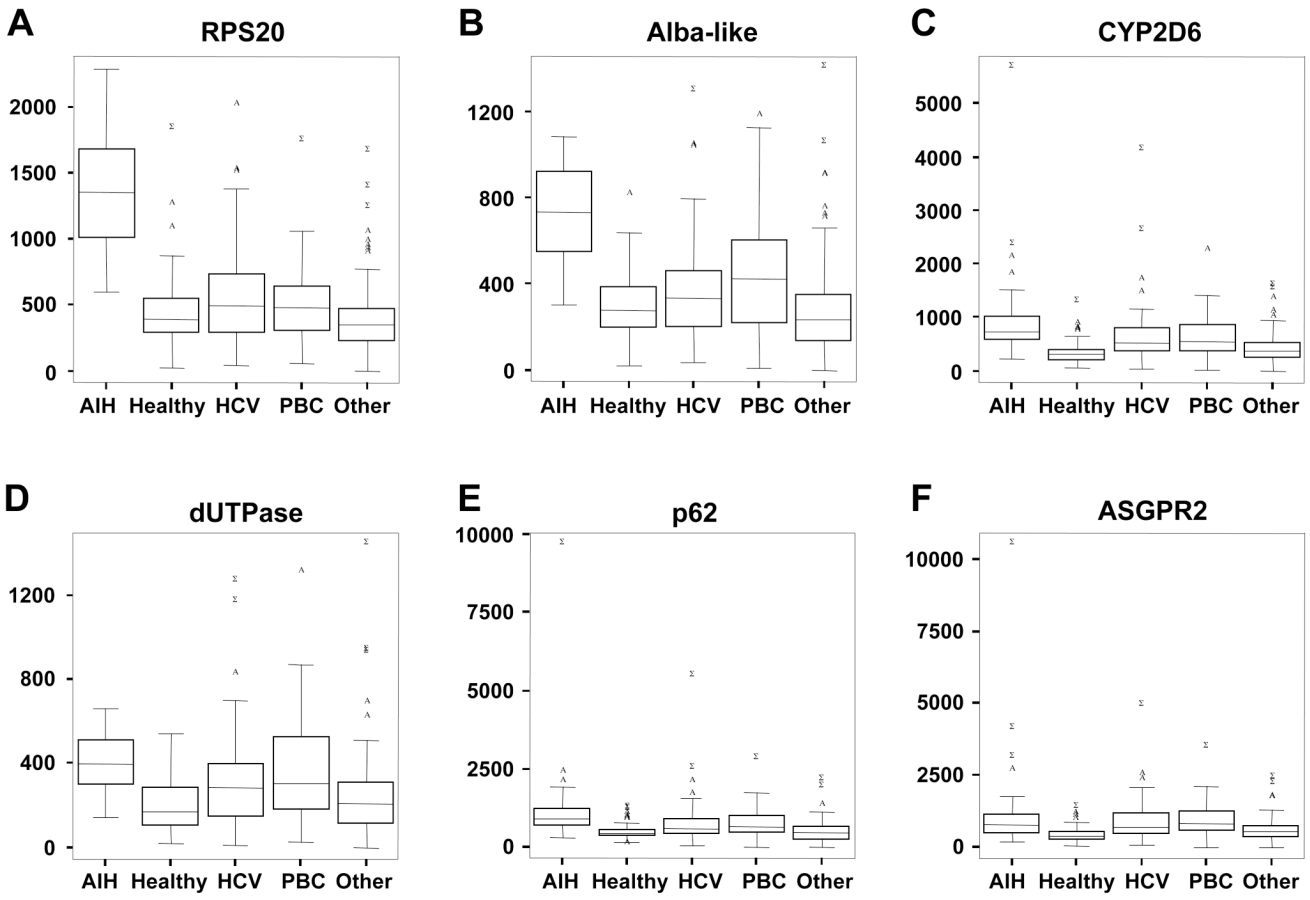

Figure 4. Box-whisker plot analysis of protein microarray probing data

All of the 278 cases were classified into the AIH, PBC, HC, healthy control, or other diseases group on the basis of clinical diagnosis. The signal distributions of each of the six autoantigens (A, RPS20; B, Alba-like; C, CYP2D6; D, dUPTase; E, p62; F, ASGPR2) reacting with the serum samples in each case are displayed. The rectangles define the interquartile range (IQR). The bar within the rectangle indicates the median value. The bars above and below the rectangles define the 1.5IQR outlier ranges. All of the extreme outliers beyond the 1.5*IQR + median are showed as black signs.

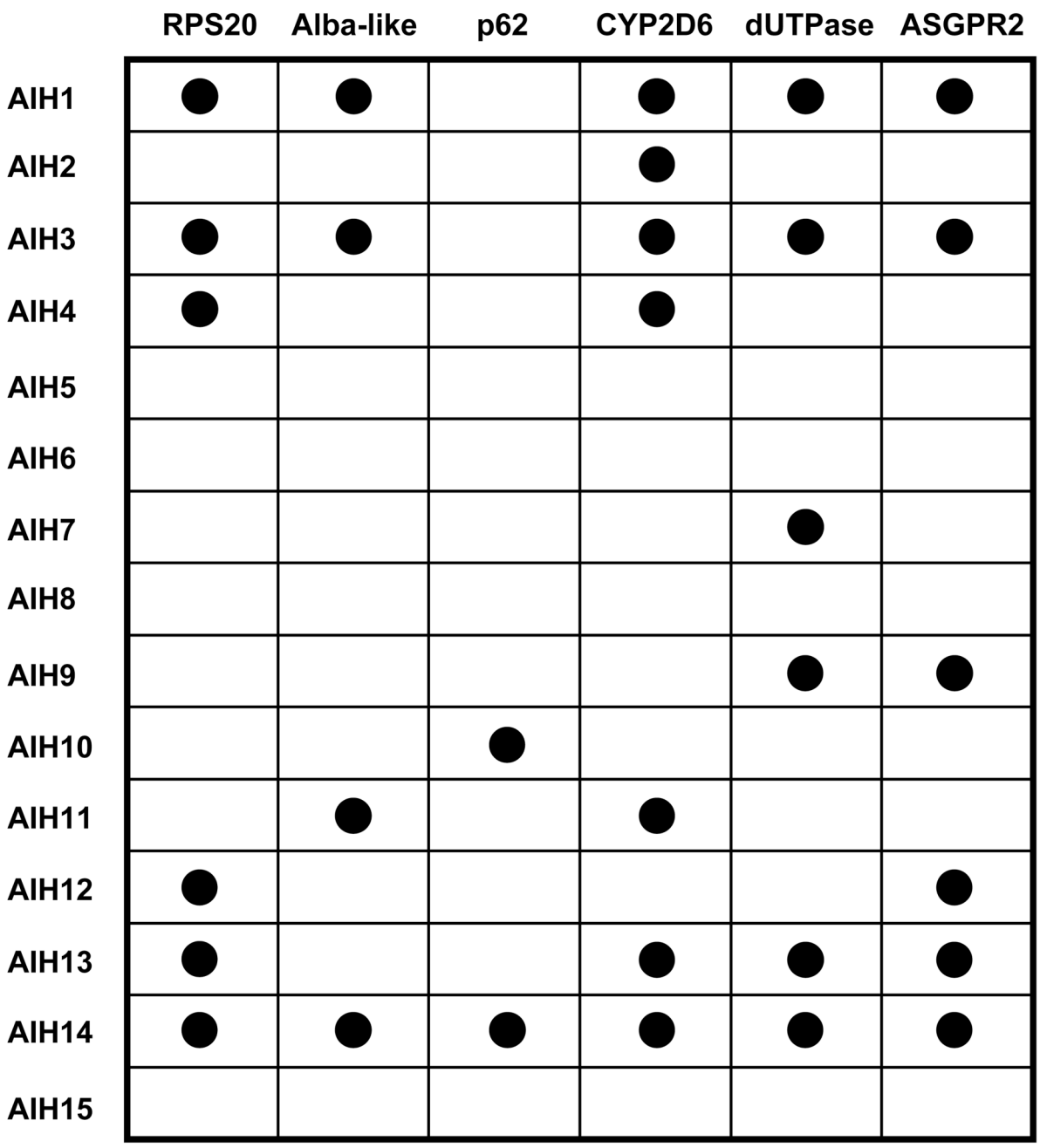

Figure 5. Comparative sensitivity of the three new and the three known autoantigens in AIH diagnosis

Proteins representing each of the six autoantigens were purified in large quantities and tested with serum samples collected from 15 AIH patients and 15 healthy subjects. Black dots indicate a positive diagnosis.

Table 1

AIH-specific autoantigen candidates identified by human protein chip serum-probing assays

Protein ID	Accession number	Positives in AIH
RPS20	Q9H6G8	15/22
Alba-like (hypothetical, Alba-like protein)	Q8N5L8	15/22
p62 (nucleoporin 62kDa)	NP_057637	14/22
CYP2D6	CAG30316	14/22
PRO0245 (dUTPase, dUTP diphosphatase)	Q9UI74	14/22
ASGPR2	P07307	13/22
GC-rich promoter binding protein 1	NP_075064	13/22
Similar to protein phosphatase 2A regulatory subunit delta isoform	Q14738	13/22
MGC16385 protein	AAH05105	13/22
Glutaredoxin-1 (thioltransferase-1, TTase-1).	P35754	13/22
LOC57002 (hypothetical protein)	NP_064577	13/22
Hypothetical protein	ENST00000333859	13/22
Ubiquitin-like protein SMT3B (HSMT3)	P55855	13/22
PRO3121	AAG35508	13/22

Table 2

Analysis of 278 sets of microarray data identifying 6 of the 14 autoantigen candidates as being specific for AIH

Candidates	Sensitivity in Groups					t-Test *p* Value	Standard Canonical Discriminant Coefficient
	AIH	Healthy	PBC	HC	Other		
RPS20	21/44	1/50	1/50	4/41	3/93	3.91E-30	0.819
	47.70%	2%	2%	9.80%	3.23%		
Alba-like	20/44	1/50	7/50	4/41	3/93	1.26E-16	0.943
	45.50%	2%	14%	9.80%	3.23%		
p62	11/44	0/50	8/50	5/41	4/93	5.43E-06	0.012
	25%	0%	16%	12.20%	4.30%		
CYP2D6	9/44	1/50	6/50	7/41	4/93	5.41E-07	0.076
	20.50%	2%	12%	17%	4.30%		
dUTPase	10/44	1/50	12/50	5/41	7/93	8.35E-04	0.112
	22.70%	2%	24%	12.20%	7.53%		
ASGPR2	9/44	1/50	14/50	10/41	5/93	5.18E-04	0.146
	20.50%	2%	28%	24.40%	5.38%		
NP_075064	0/44	0/50	8/50	4/41	6/93	0.720	−0.235
	0%	0%	16%	9.80%	6.45%		
Q14738	0/44	0/50	8/50	5/41	9/93	0.247	−0.604
	0%	0%	16%	12.20%	9.68%		
MGC16385	1/44	0/50	13/50	4/41	6/93	0.872	−0.481
	2.30%	0%	26%	9.80%	6.45%		
Glutaredoxin-1	3/44	1/50	8/50	7/41	6/93	0.731	−0.115
	6.8%	2%	16%	17%	6.45%		
LOC57002	1/44	0/50	12/50	4/41	5/93	0.420	−0.152
	2.3%	0%	24%	9.8%	5.38%		
ENST00000333859	0/44	0/50	6/50	4/41	8/93	0.075	−0.202
	0%	0%	12%	9.8%	8.60%		
SMT3B	2/44	0/50	11/50	1/41	3/93	0.488	0.136

Candidates	Sensitivity in Groups					t-Test *p* Value	Standard Canonical Discriminant Coefficient
	AIH	Healthy	PBC	HC	Other		
	4.5%	0%	22%	2.4%	3.23%		
PRO3121	0/44	0/50	11/50	5/41	4/93	0.531	−0.120
	0%	0%	22%	12.2%	4.30%		

Table 3

Prediction performance of two classification methods for the training cases

Prediction method	Total Cases	Correctly Classified Cases	False-positive	False-negative	Sensitivity[a]	Specificity[b]	Accuracy
Logistic Regression	278	255	9	14	68.2%	96.2%	91.7%
Discriminant Analysis	278	255	16	7	84.1%	93.2%	91.7%

[a] Sensitivity is defined as the true positive rate.

[b] Specificity is defined as true negative rate.

Table 4

Prediction performance of the two classification methods in a double-blind experiment

Prediction method	Total Cases	Correctly Classified Cases	False-positive	False-negative	Sensitivity[a]	Specificity[b]	Accuracy
Logistic Regression	41	35	0	6	60%	100%	85.4%
Discriminant Analysis	41	36	0	5	66.7%	100%	87.8%

[a] Sensitivity is defined as the true positive rate.

[b] Specificity is defined as true negative rate.