Antibody Profiling of Kawasaki Disease Using *Escherichia coli* Proteome

Microarrays

Ho-Chang Kuo¹, Ying-Hsien Huang¹, Feng-Hsiang Chung²,³, Po-Chung Chen²,³, Tzu-Cheng Sung²,³, Yi-Wen Chen²,³, Kai-Sheng Hsieh¹, Chien-Sheng Chen²,³,⁴ CA, Guan-Da Syu²,³ CA

¹Department of Pediatrics and Kawasaki Disease Center, Kaohsiung Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Kaohsiung, Taiwan 83301.

²Graduate Institute of Systems Biology and Bioinformatics, National Central University, Taoyuan, Taiwan 32001.

³Department of Biomedical Sciences and Engineering, National Central University, Taoyuan, Taiwan 32001.

⁴Department of Food Safety / Hygiene and Risk Management, National Cheng Kung University, Tainan, Taiwan 70101

CA Corresponding Authors:

Guan-Da Syu, Ph.D.

Department of Pharmacology and Molecular Sciences,
Johns Hopkins School of Medicine

Edward D. Miller Research Building, 733 N. Broadway, Baltimore, MD 21205

Tel: 410-502-2807; Fax: 410-502-1872

E-mail: gsyu1@jhmi.edu

Chien-Sheng Chen, Ph.D.

Graduate Institute of Systems Biology and Bioinformatics

Department of Biomedical Science and Engineering

National Central University

No. 300, Rd. Zhongda, Dist. Zhongli

Taoyuan City 32001, Taiwan, ROC.

Tel: 886-3-4227151 ext. 36103; Fax: 886-3-4253427

E-mail: cchen103@gmail.com

Running title: Antibody Profiling of Kawasaki Disease

Abbreviations: Kawasaki disease (KD), intravenous immunoglobulin (IVIG),
non-fever control (NC), fever control (FC), KD before intravenous immunoglobulin
treatment (KD1), KD at least 3 weeks after treatment (KD3).
Summary

Kawasaki disease (KD) is a form of systemic vasculitis that generally occurs in children under 5 years old. Currently, KD is still diagnosed according to its clinical symptoms, including prolonged fever, skin rash, conjunctivitis, neck lymphadenopathy, palm erythema, and oral mucosa changes. Since KD is a type of inflammation without specific marker for diagnosis, we plan to profile the plasma antibodies by using *E. coli* proteome microarray and analyze the differences between KD and healthy subjects. Plasmas were collected from KD patient before intravenous immunoglobulin treatment (KD1), at least 3 weeks after treatment (KD3), non-fever control (NC), and fever control (FC) children. The initial screening, which consisted of 20 KD1, 20 KD3, 20 NC, and 20 FC, were explored using *E. coli* proteome microarrays (~4200 unique proteins). About ~70 proteins were shown to have high accuracy, e.g., 0.78~0.92, with regard to separating KD1, KD3, NC, and FC. Those proteins were then purified to fabricate KD focus arrays for training (N = 20 each) and blind-testing (N = 20 each). It only took 125 pl of plasma, less than a drop of blood, in the focus array assays. The AUC scores for blind tests of KD1 vs. NC (17 protein markers), KD1 vs. FC (20 protein markers), KD3 vs. NC (9 protein markers), and KD1 vs. KD3 (6 protein markers) were 0.84, 0.75, 0.99 and 0.98, respectively.
This study is the first to profile plasma antibodies in KD and demonstrate that an *E. coli* proteome microarray can screen differences among patients with KD, non-fever controls, and fever controls.

**Keywords**: Antibody, Biomarker, High throughput, Kawasaki disease, Proteome microarray
**Introduction**

Kawasaki disease (KD) is a form of acute systemic vasculitis syndrome with an unknown etiology that most commonly occurs in children under the age of 5 years old. KD is characterized by prolonged fever, bilateral conjunctivitis, diffuse mucosal inflammation/strawberry tongue, polymorphous skin rashes, indurative edema of the hands and feet associated with peeling of the finger tips, and non-suppurative neck lymphadenopathy (1). The most serious complication of KD is the formation of coronary artery lesions. A sequela of the vasculitis with coronary artery aneurysm has developed in 20-25% of untreated children, but a single high-dose of intravenous immunoglobulin (IVIG) can lower the incidence of aneurysm to 3-5% (2).

Physicians diagnose KD according to fever days and five main clinical symptoms, but this method may not be objective enough. Although some lab tests such as C-reactive protein and erythrocyte sedimentation rate may help to confirm the inflammation (3), they are not specific markers for KD. Till now, there is no biochemistry diagnosis kit available for KD. It has been reported that KD patients have antibody differences in their blood, such as anti-endothelial cell antibodies and anti-neutrophil antibodies (4-6). However, other groups do not find these antibody differences (7). This controversial result indicates the need of high throughput techniques for screening antibody markers.
More than 85% of KD occurs in children under the age of 5 years. In infant or early childhood, gut microbiota plays an important role in immune homeostasis and autoimmunity (8-10). Since *E. coli* is a type of common commensal intestinal bacteria in humans, it is revealed to be one of the first foreign substances that a newborn encounters. Therefore, *E. coli* logically contributes to humoral immune responses in infants and may be helpful with regard to identifying KD disease markers. We use high-throughput techniques to express and purify ~4200 *E. coli* proteins and develop *E. coli* proteome microarrays (11-14). These *E. coli* proteome microarrays have been successfully profiling antibodies in the bloodstream and identifying antibody biomarkers in inflammatory bowel disease (12) and bipolar disorder (14). Since most KD symptoms have been correlated to immune dysfunction, we hypothesize that antibodies in the plasma behave differently in KD patients. Furthermore, the antibodies may be altered in KD patients after treatment and recovery. To investigate our hypothesis, we measured the antibody amount in the plasma and profiled the antibodies using thousands of *E. coli* antigens. In this study, we found different antibody profiles in KD patients and we confirmed some antibodies with high accuracy in the single-blind test.
Materials and Methods

Patients and controls

In this study, the KD subjects (n = 60) consisted of children who met the criteria for KD (15) and were treated with IVIG at Kaohsiung Chang Gung Hospital. The patients were initially treated with a single dose of IVIG (2 g/kg) during a 12-hour period as in our previous reports. Blood samples were obtained prior to IVIG treatment (pre-IVIG, KD1), within 3 days following IVIG treatment (post-IVIG, KD2) as acute stage samples, and at least 3 weeks following IVIG (KD3). We excluded patients whose symptoms did not match the KD criteria, had an acute fever for less than 5 days, or had an incomplete collection of pre- and post-IVIG blood samples. Age-matched febrile patients (n = 60) and non-fever children (n = 60) from outpatient clinic were included as controls. Febrile control patients consisted of children admitted to the ward with an upper and/or lower respiratory tract infection. Blood samples were collected in 10 mL heparin tubes (BD, catalog No. 367874), kept in room temperature, and harvested plasma by centrifuge at 3000 rpm for 10 minutes within 24 hours. Plasma samples were aliquoted into several copies and stored immediately in -80 °C refrigerator until detection. The Institutional Review Board of Kaohsiung Chang Gung Memorial Hospital (IRB No. 102-1015A3) approved this
study. Informed consent was obtained from all individuals’ parents or guardians prior to carrying out the study.

Measuring plasma IgG

We determined plasma IgG concentrations using the anti-human IgG ELISA kit (eBioscience) in accordance with the manufacturer’s instructions. In short, 96-well plates were coated with a capture antibody, washed, blocked, and incubated with 750,000 fold plasma diluents. After being washed several times, we applied the detection antibody and incubated it for another hour with shaking. After another several washes, we sequentially added tetramethylbenzidine, stopped the solution, and acquired data using the ELISA reader with O.D. 450 nm (BioTek, synergy 2).

Creating E. coli proteome microarrays

We modified the methods for high-throughput protein expression, purification, and printing from our previous study (11). In brief, His-tag proteins were overexpressed in an E. coli ASKA collection and purified with resin beads in a high-throughput manner. To create the E. coli proteome microarrays, ~4200 purified proteins were identified on each aldehyde slide in duplicate using SmartArrayer 136 (CapitalBio). Protein microarrays were immobilized at 4 °C for 8 h and then stored at
Antibody profiling using E. coli proteome chips

To use, the E. coli proteome chips were thawed in TBST and blocked with 3% BSA for 1 hour. Plasmas were diluted to 1.25 mg/dl IgG in 1% BSA of TBST as predetermined by the ELISA kit. The blocked proteome chips were incubated with plasma dilutant for 1 hour and shaken at 50 rpm. After washing 3 x 10 min with TBST, the chips were incubated with an anti-human antibody cocktail, which consisted of Dylight550 labeled anti-human IgM antibody (Bethyl Laboratories) and Dylight650 labeled anti-human IgG antibody (Bethyl Laboratories), for 1 hour. Each chip was washed 3 x 10 min with TBST and scanned using a chip scanner (CapitalBio, LuxScan™).

Data processing and analysis

Signal intensity for each spot is extracted using GenePix Pro 6.0 software. We averaged the duplicates in each chip and then normalized all data by mean scaling. In order to maintain data quality, we applied principal component analysis to remove poorly performing chips. We further used supervised machine learning (16) to select a small set of markers; our selection process included significance analysis of
microarrays, support vector machine, and 10-fold cross-validation. K-means clustering (17) was used to demonstrate the differentiation power of the selected markers. Finally, we calculated accuracy using the sum of true positive and true negative divided by the sum of positive and negative.

Creating focus array and plasma assays

The selected 73 proteins were purified and printed on the aldehyde slides (CapitalBio) using the same protocol that was used for the *E. coli* proteome microarray. The format was designed to fit a 14-well cassette (Figure 3). Each protein was printed in duplicate at 4 °C and then immobilized for 8 hours. Focus array chips were kept at -80 °C until they were needed to carry out the experiments.

Additional plasma of NC, KD1, KD3, and FC (n = 20) were probed using focus arrays. Focus arrays were blocked using 3% BSA for 1 hour. After blocking, the chips were washed with TBST for 10 min and assembled into 2 x 7 well cassettes (Arrayit). The plasmas were diluted into 1:800 with 1% BSA of TBST, and 100 μl was placed into each well. The chips were incubated for 1 hour with shaking, disassembled, and washed for another 3 x 10 min with TBST. The detection antibodies, washing steps, and scanning protocols were as same as described in the previous section.
Training and single-blind testing

To prevent potential bias, different personnel performed the plasma coding, plasma assay, and data analysis. Another set of plasmas were randomly coded with X01-X80. Signal intensities were extracted by GenePix Pro 6.0 in the foreground minus the background format. Once the duplicates were averaged, each block was normalized by mean scaling, and then the negative values were replaced with 1. With regard to the training section, all the plasmas were labeled with NC, KD1, KD3, or FC (n=20). We applied a logistical regression model (LRM) to systematically combine different biomarkers, including IgG and IgM responsive proteins with different signal scales in the training group, and then evaluated the performance of these protein groups by examining the area under the curve (AUC). We conducted LRM analysis using the “glm” of the package “stat” and calculated the AUC score with the “performance” function from the package “ROCR” in the R environment (R version 3.0.3). In single-blind testing, we coded all the plasmas with random numbers and compared the results with the information provided by the pediatric doctors. The AUC score cutoff for the training group was controlled by 0.9, and the protein groups would be formed as a unique factor. We further examined the performance of the testing group, for which 20 additional patients were chosen. Evaluation of the performance of both the training and testing groups was performed by comparing
accuracy, specificity, sensitivity, and AUC scores.
Results

Physiological parameters of KD and control subjects

This study consisted of 20 KD1, 20 KD3, 20 FC, and 20 NC in order to determine plasma IgG levels and further investigate using *E. coli* proteome chips. Another 20 KD1, 20 KD3, 20 FC, and 20 NC were collected to serve as focus arrays, while an additional 20 KD1, 20 KD3, 20 FC, and 20 NC were enrolled for blind testing. Most of the subjects enrolled in this study were under 5 years old and were compatible with the predominant onset age of Kawasaki disease. No significant difference with regard to gender distribution was found between the groups (Supplementary Table 1).

Antibody profiling of KD and control subjects using *E. coli* proteome chips

Before carrying out the antibody profiling, we determined the IgG concentrations in the plasmas. Surprisingly, we found some disease-related differences in the IgG concentrations (Supplementary Figure 1). Reduced IgG levels were only observed in KD1 but not NC or FC. These findings indicated that plasma IgG plays a crucial role in KD, thus giving us a strong reason to profile the antibodies.

We used high-density *E. coli* proteome chips, which consisted of ~4200 proteins,
to profile the antibodies in human plasmas (Figure 1). We performed serial optimization of plasma dilutions and detection antibody concentrations in order to maximize the detection signal while minimizing the nonspecific bindings. To ensure equivalent comparisons, a fixed IgG concentration (1.25 mg/dl IgG, about 1000-fold dilution) was used to dilute the plasma samples. Each plasma sample was probed with an E. coli proteome chip. We used a total of 4 x 20 proteome chips to assay NC, FC, KD1, and KD3 (each groups N = 20). After being incubated with chips, the plasma antibodies were visualized using Dylight650 labeled anti-IgG and Dylight550 labeled anti-IgM antibodies. Anti-IgG and anti-IgM were scanned using different channels and combined into a dual-color chip image (Figure 1B).

*Selecting markers for KD and control subjects from antibody profiling*

To identify markers for healthy individuals and KD sufferers, we normalized signals from each chip with mean scaling and sent them for supervised machine learning (Figure 1C), which included three steps: selecting significant hits, supervising learning, and performing cross-validation. We selected a small number of markers based on these strategies, which are shown in the heat map (Figure 2). In Comparing the KD1 group to the NC group, we found 6 IgG markers with sensitivity 68%, specificity 94% and accuracy 81%, and 11 IgM markers with sensitivity 77%,
specificity 94% and accuracy 84% (Figure 2A). In comparing the KD1 group to the FC group, we found 12 IgG markers with sensitivity 79%, specificity 90% and accuracy 84%, and 13 IgM markers with sensitivity 74%, specificity 90% and accuracy 82% (Figure 2B). In comparing the KD3 group to the NC group, we found 11 IgG markers with sensitivity 84%, specificity 100% and accuracy 92%, and 9 IgM markers with sensitivity 78%, specificity 100% and accuracy 89% (Figure 2C). In comparing the KD1 group to the KD3 group, we found 16 IgG markers with sensitivity 84%, specificity 84% and accuracy 84%, and 11 IgM markers with sensitivity 94%, specificity 63% and accuracy 78% (Figure 2D). In general, the markers indicating KD disease and those indicating KD recovery both demonstrated remarkable accuracy.

Designing and creating focus arrays

Of the ~4200 different *E. coli* proteins, we successfully identified ~70 promising diagnosis and recovery markers. To further support their clinical value, we created focus arrays containing all of the IgG and IgM markers. We printed each marker in duplicate in each 75 mm x 25 mm slide and formed 14 identical blocks (Figure 3). To imitate clinical use, we did not adjust the dilution factor based on the IgG concentration, but instead set the dilution factor to 800-fold, which provided the best
signal. This assay only took 125 pl of plasma which is less than a drop of blood. We used an additional \( n = 20 \) non-coded plasmas in the training section and \( n = 20 \) coded plasmas for the single-blind testing (Figure 1C).

**Training and single-blind testing**

The receiver operating characteristic (ROC) curve can provide comprehensive analysis for summarizing the accuracies of predictions (18, 19); therefore, we applied the ROC curve to compare the different tests of the protein biomarkers discovered from the supervised machine learning results. Furthermore, we performed logistic regression model (LRM) analysis to combine these protein biomarkers to determine whether these combinations would increase overall performance. Since each point on the ROC curve represents the true positive rate (sensitivity) and false positive rate (1 - specificity) of a particular cut-off value, we can consider the area under the curve (AUC) as an overall performance indicator, while an AUC score close to 1 indicates an excellent diagnostic test, and an AUC value close to 0.5 indicates a poor diagnostic test. We first researched the individual performance of selected proteins in the training set of KD1 vs. NC and found that AUC scores for the selected proteins ranged from 0.68 to 0.77 (the colored solid lines showed in the left of Figure 3B, Supplementary Table 3). After performing LRM analysis on these protein biomarkers, we created a
combinational protein group using 17 IgM-responsive proteins, which increased the AUC score for the protein group to 0.91 (the black solid line in the left of Figure 3B). The protein group also showed good performance on the single-blind test using an additional testing set of KD1 vs. NC, in which the AUC score reached 0.84 (the red solid line shown in the right of Figure 3B). Furthermore, most of the proteins in this group showed significant immunogenic responses in the KD1 group compared with the NC group in the single-blind test (Figure 3C, 14 of 17 proteins showed significant differential bindings between KD1 patients and NC people with a \( p \)-value smaller than 0.05). We also investigated conservative motifs for these IgM-responsive proteins and found a unique sequence named [EI]IDALVEK[QA][IL][EN][ERD]LI (Supplementary Figure 2). These proteins are also involved in nucleotide binding and are extrinsic to membrane and pentose-glucuronate metabolism. Such results suggest that these IgM-responsive proteins may contribute to Kawasaki disease and that the consensus motif may provide predictive values for further testing.

Since children with Kawasaki disease have fevers, we also included children with fever symptoms but that were not diagnosed with Kawasaki disease as our other group of negative controls. We selected 20 proteins using the training set; the AUC scores ranged from 0.56 to 0.63 (the colored solid lines in the left of Figure 4A), and the LRM results for these proteins increased the AUC score to 0.82 in the training set.
of the KD1 group vs. the FC group (the black solid line in the left of Figure 4A, Supplementary Table 3). We further examined predictive power using this protein group in the testing set, and the AUC score was only 0.75 for this single-blind test (the red solid line in the right of Figure 4A). However, the boxplot demonstrated that none of the proteins significantly differentiated KD1 patients from FC controls (Figure 4B, \(p\)-values greater than 0.05).

Regarding Kawasaki disease patients at least 3 weeks after IVIG treatment (KD3), we are interested in whether any prognostic biomarker could differentiate KD3 from normal controls (NC) or patients prior to receiving IVIG treatment (KD1). The ROC curve results revealed excellent performance in the training set of the KD3 group vs. the NC group, and the AUC scores for these nine proteins ranged from 0.93 to 0.96 (the colored solid lines in the left of Figure 5A, Supplementary Table 3). After LRM analysis, the overall performance improved, and the AUC score was 0.98 for this protein group (the black solid line in the left of Figure 5A, Supplementary Table 3). In the single-blind test, this protein group continues to have predictive power, and the AUC score is 0.99 for the testing set (the red solid line in the right of Figure 5A). All nine of these proteins significant immunogenic responses against IgG antibodies in KD3 patients compared with the NC group (Figure 5B, \(p\)-values are less than 0.001).
Regarding the KD1 group vs the KD3 group, we discovered some proteins that showed excellent prognostic values in transitioning from KD1 to KD3. Six proteins, were selected using the training set of KD1 vs KD3 groups, demonstrated very good performance on the ROC curve, the AUC scores ranged from 0.93 to 0.95, and the score for this protein group was 0.97 using LRM analysis (the black solid line showed the group result in the left of Figure 6A, Supplementary Table 3). In the single-blind test, the combination of this protein group also showed very good performance, resulting in an AUC score of 0.98 when using the additional testing set (the red solid line in the right of Figure 6A). Like in the KD3 vs NC groups, these six proteins were also IgG-responsive proteins and demonstrated significant immunogenic responses in KD3 patients when compared to the KD1 patients (Figure 6B, all $p$-values are less than 0.001).
Discussion

Kawasaki disease is a form of acute febrile vasculitis that most commonly affects children under the age of 5 years and can lead to multiple organ injuries, including of the kidneys and heart. One common feature of KD patients is pyuria, which affects 30-80% (20). Sterile pyuria most commonly occurs in KD patients ≤ 1 year of age. Although not always sterile in KD patients, pyuria can be the result of a urinary tract infection. *E. coli* and *Klebsiella oxytoca* have previously been reported as causative pathogens for Kawasaki disease (20, 21). The clinical phenotypes do not differ between patients with and those without urinary tract infection. These evidence indicates that *E. coli* may be a relevant pathogen of KD.

In this study, we analyzed antibody profiles of KD1 and KD3 serum and compared them with FC and NC. We found that KD3 antibody profiles are very different from those of NC and KD1, indicating that a huge change occurs with Kawasaki development or IVIG treatment. In general, the half-life of IVIG in blood is 21 days (22), which may partially explain the antibody profile difference. Another possible reason is adaptive immunity, which would produce IgG, in which case the peak is 21 days (23). Both IVIG and adaptive immunity would contribute to most, if not all, of the dramatic changes in the antibody profiles. Interestingly, all six classifying markers for KD3 vs KD1 are also good classifiers for KD3 vs NC, which
suggests that the KD3-developed antibodies against these E.coli proteins are so
unique that neither KD1 nor NC subjects have them, thus indicating that adaptive
immunity may take part in developing some unique antibodies from KD1 to KD3.

Of all the identified classifiers for KD1 vs. NC, KD1 vs. FC, KD3 vs. NC, and
KD3 vs. KD1, we found that the znuC protein is the most common protein, which
indicates that znuC is clearly an antigen from KD1 to KD3. Interestingly, class
differences appear in the anti-znuC antibodies, e.g., anti-znuC IgM in KD1 and IgG in
KD3. ZnuC is an ATP-binding protein and a part of the ABC transporter for zinc. It is
located in the bacterial surface and is vital for the survival of stress-induced bacteria
(24). On one hand, znuC is conserved in many bacterial species, such as E.coli,
Klebsiella oxytoca, Shigella flexneri, Salmonella enterica, and Pasteurella multocida.
Both E. coli and Klebsiella oxytoca infections have been reported in KD (20, 21). On
the other hand, zinc is important for cardiovascular health in humans (25).
Cardiovascular inflammation is the most severe symptom of KD patients (2).
Therefore, znuC is an ideal target for KD patients to develop a specific antibody, thus
making it a perfect marker for KD diagnosis.

Our blind test results revealed that our identified classifying protein markers
could effectively classify KD1 vs NC, KD1 vs FC, KD3 vs NC, and KD3 vs KD1
with an AUC value of at least 0.75 (Supplementary Table 2). Of those, KD3 patients
could be easily distinguished from KD1 and NC. However, KD1 cannot be well
separated from FC, which may be due to the complexity of the fever controls. Further
studies are needed to separate FC subtypes and compared to KD1. This study
combines proteome microarray screening, bioinformatics analysis, and validation to
identify reliable markers for KD diagnosis. We only required a drop of blood for our
specimens. These platforms and markers may improve the understanding of the
diagnosis and prognosis of KD.
References

1. Kuo, H. C., Yang, K. D., Chang, W. C., Ger, L. P., and Hsieh, K. S. (2012) Kawasaki disease: an update on diagnosis and treatment. Pediatr Neonatol 53, 4-11

2. Burns, J. C., and Glode, M. P. (2004) Kawasaki syndrome. Lancet 364, 533-544

3. Newburger, J. W., Takahashi, M., Gerber, M. A., Gewitz, M. H., Tani, L. Y., Burns, J. C., Shulman, S. T., Bolger, A. F., Ferrieri, P., Baltimore, R. S., Wilson, W. R., Baddour, L. M., Levison, M. E., Pallasch, T. J., Falace, D. A., and Taubert, K. A. (2004) Diagnosis, treatment, and long-term management of Kawasaki disease: a statement for health professionals from the Committee on Rheumatic Fever, Endocarditis, and Kawasaki Disease, Council on Cardiovascular Disease in the Young, American Heart Association. Pediatrics 114, 1708-1733

4. Fujieda, M., Oishi, N., and Kurashige, T. (1997) Antibodies to endothelial cells in Kawasaki disease lyse endothelial cells without cytokine pretreatment. Clinical and experimental immunology 107, 120-126

5. Chan, T. M., Frampton, G., Jayne, D. R., Perry, G. J., Lockwood, C. M., and Cameron, J. S. (1993) Clinical significance of anti-endothelial cell antibodies in systemic vasculitis: a longitudinal study comparing anti-endothelial cell antibodies and anti-neutrophil cytoplasm antibodies. American journal of kidney diseases : the
6. Frampton, G., Jayne, D. R., Perry, G. J., Lockwood, C. M., and Cameron, J. S. (1990) Autoantibodies to endothelial cells and neutrophil cytoplasmic antigens in systemic vasculitis. *Clinical and experimental immunology* 82, 227-232

7. Nash, M. C., Shah, V., Reader, J. A., and Dillon, M. J. (1995) Anti-neutrophil cytoplasmic antibodies and anti-endothelial cell antibodies are not increased in Kawasaki disease. *British journal of rheumatology* 34, 882-887

8. Lundell, A. C., Bjornsson, V., Ljung, A., Ceder, M., Johansen, S., Lindhagen, G., Tornhage, C. J., Adlerberth, I., Wold, A. E., and Rudin, A. (2012) Infant B cell memory differentiation and early gut bacterial colonization. *Journal of immunology* 188, 4315-4322

9. Cenit, M. C., Matzaraki, V., Tigchelaar, E. F., and Zhernakova, A. (2014) Rapidly expanding knowledge on the role of the gut microbiome in health and disease. *Biochimica et biophysica acta* 1842, 1981-1992

10. Wu, H. J., and Wu, E. (2012) The role of gut microbiota in immune homeostasis and autoimmunity. *Gut microbes* 3, 4-14

11. Chen, C. S., Korobkova, E., Chen, H., Zhu, J., Jian, X., Tao, S. C., He, C., and Zhu, H. (2008) A proteome chip approach reveals new DNA damage recognition activities in Escherichia coli. *Nature methods* 5, 69-74
12. Chen, C. S., Sullivan, S., Anderson, T., Tan, A. C., Alex, P. J., Brant, S. R.,
Cuffari, C., Bayless, T. M., Talor, M. V., Burek, C. L., Wang, H., Li, R., Datta, L. W.,
Wu, Y., Winslow, R. L., Zhu, H., and Li, X. (2009) Identification of novel serological
biomarkers for inflammatory bowel disease using Escherichia coli proteome chip.
Molecular & cellular proteomics : MCP 8, 1765-1776

13. Ho, Y. H., Sung, T. C., and Chen, C. S. (2012) Lactoferricin B inhibits the
phosphorylation of the two-component system response regulators BasR and CreB.
Molecular & cellular proteomics : MCP 11, M111 014720

14. Chen, P. C., Syu, G. D., Chung, K. H., Ho, Y. H., Chung, F. H., Chen, P. H., Lin,
J. M., Chen, Y. W., Tsai, S. Y., and Chen, C. S. (2015) Antibody profiling of bipolar
disorder using Escherichia coli proteome microarrays. Molecular & cellular
proteomics : MCP 14, 510-518

15. Newburger, J. W., Takahashi, M., Gerber, M. A., Gewitz, M. H., Tani, L. Y.,
Burns, J. C., Shulman, S. T., Bolger, A. F., Ferrieri, P., Baltimore, R. S., Wilson, W. R.,
Baddour, L. M., Levison, M. E., Pallasch, T. J., Falace, D. A., and Taubert, K. A.
(2004) Diagnosis, treatment, and long-term management of Kawasaki disease: a
statement for health professionals from the Committee on Rheumatic Fever,
Endocarditis and Kawasaki Disease, Council on Cardiovascular Disease in the Young,
American Heart Association. Circulation 110, 2747-2771
16. Glaab, E., Garibaldi, J. M., and Krasnogor, N. (2009) ArrayMining: a modular web-application for microarray analysis combining ensemble and consensus methods with cross-study normalization. *BMC bioinformatics* 10, 358

17. de Hoon, M. J., Imoto, S., Nolan, J., and Miyano, S. (2004) Open source clustering software. *Bioinformatics* 20, 1453-1454

18. Warnock, D. G., and Peck, C. C. (2010) A roadmap for biomarker qualification. *Nature biotechnology* 28, 444-445

19. Mandrekar, J. N. (2010) Receiver operating characteristic curve in diagnostic test assessment. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer* 5, 1315-1316

20. Watanabe, T. (2015) Pyuria in patients with Kawasaki disease. *World journal of clinical pediatrics* 4, 25-29

21. Husain, E. H., and Al-Rashid, M. (2011) Kawasaki disease in association with urinary tract infection. *Indian pediatrics* 48, 808-809

22. Bonilla, F. A. (2008) Pharmacokinetics of immunoglobulin administered via intravenous or subcutaneous routes. *Immunology and allergy clinics of North America* 28, 803-819, ix

23. Miao, H., Hollenbaugh, J. A., Zand, M. S., Holden-Wiltse, J., Mosmann, T. R., Perelson, A. S., Wu, H., and Topham, D. J. (2010) Quantifying the early immune
response and adaptive immune response kinetics in mice infected with influenza A virus. *Journal of virology* 84, 6687-6698

24. Al Mamun, A. A., Lombardo, M. J., Shee, C., Lisewski, A. M., Gonzalez, C., Lin, D., Nehring, R. B., Saint-Ruf, C., Gibson, J. L., Frisch, R. L., Lichtarge, O., Hastings, P. J., and Rosenberg, S. M. (2012) Identity and function of a large gene network underlying mutagenic repair of DNA breaks. *Science* 338, 1344-1348

25. Little, P. J., Bhattacharya, R., Moreyra, A. E., and Korichneva, I. L. (2010) Zinc and cardiovascular disease. *Nutrition* 26, 1050-1057
Footnotes

Sources of Funding

This study was supported by grants from the Ministry of Science and Technology, Taiwan (MOST 105-2314-B-182-050-MY3, MOST 104-2320-B-008-002-MY3, and MOST 104-2627-M-008-001), the Aim for the Top University Project (104G705-4), the Ministry of Health and Welfare (PMRPG8E0011) and the Chang Gung Memorial Hospital (CMRPG8C1082, CMRPG8B0212, CMRPG8D1562, CORPG8F0011, CMRPG8E1611, and CMRPG8D0521 to HC-K) of Taiwan.

Conflict of Interest Statement

All of the authors declare that they have no financial relationships and no conflicts of interest to disclose with regard to this article.
Figure Legends

Figure 1. Profiling antibody responses in KD patients and controls using *E. coli* proteome chips. (A) Plasmas from normal or KD patients were collected and incubated with *E. coli* proteome chips. The IgG and IgM from the plasmas were visualized using Cy3-labeled anti-IgM and Cy5-labeled anti-IgG, respectively. (B) Typical image of profiling plasma IgG and IgM using an *E. coli* proteome chip. Approximately 4200 proteins were on a chip; the green signal represented Cy3-labeled anti-IgM and the red label represented Cy5-labeled anti-IgG. (C) Plasmas from NC, KD1, KD3, and FC were collected as previously described in the materials and methods section (n = 20). Plasma IgG concentrations were determined using the ELISA kit and adjusted to 1.25 mg/dl prior to the chip assays. After normalization, significantly changed markers were identified by supervised machine learning. The KD focus array was created according to those markers and further validated by training (n = 20) and testing (n = 20) with other plasma samples.

Figure 2. Novel KD markers selected from antibody profiling using *E. coli* proteome chips. We identified disease relevant markers using supervised machine learning. This method included significance analysis of microarrays, support vector machine, and
10-fold cross-validation. Results were demonstrated by k-means clustering and shown with a heat-map. (A) KD1 vs. NC markers. (B) KD1 vs. FC markers. (C) KD3 vs. NC markers. (D) KD1 vs. KD3 markers.

Figure 3. Focus array design and training/testing for KD1 vs. NC.

(A) The protein markers shown in Figure 2 were purified and printed on aldehyde slides in a duplicate manner in 14 identical blocks. Plasma samples for validation were incubated using 14-well chambers. The top 12 wells were incubated with plasmas. The bottom 2 wells were incubated with Cy5-labeled anti-his to serve as a quality reference. After incubating with plasmas or anti-his, incubation chambers were disassembled and incubated with Cy3-labeled anti-IgM and Cy5-labeled anti-IgG. (B) Training and testing results for the KD1 group vs the NC group. In the training section, n = 20 samples were used to select the best group based on the logistic regression model. The same group was used for n = 20 single-blind testing. The decision power was defined as the area under the curve (AUC). (C) The signal intensities of the committee markers in the single-blind test. Data are presented as mean± SEM and were analyzed by student t-test, * p < 0.05, ** p < 0.01, and *** p < 0.001.
Figure 4. Training/testing for KD1 vs. FC.

(A) Training and testing results for the KD1 group vs the FC group. (B) The signal intensities of KD1 vs FC markers in the single-blind test. Data are presented as mean±SEM and were analyzed by student t-test, * p < 0.05, ** p < 0.01, and *** p < 0.001.

Figure 5. Training/testing for KD3 vs. NC.

(A) Training and testing results for the KD3 group vs the NC group. (B) The signal intensities of KD3 vs NC markers in the single-blind test. Data are presented as mean±SEM and were analyzed by student t-test, * p < 0.05, ** p < 0.01, and *** p < 0.001.

Figure 6. Training/testing for KD1 vs. KD3.

(A) Training and testing results for the KD1 group vs the KD3 group. (B) The signal intensities of KD1 vs KD3 markers in the single-blind test. Data are presented as mean±SEM and were analyzed by student t-test, * p < 0.05, ** p < 0.01, and *** p < 0.001.
Figure 1.
Figure 2.

A. KD1 vs. NC, IgG, accuracy = 0.81

B. KD1 vs. FC, IgG, accuracy = 0.84

C. KD3 vs. NC, IgG, accuracy = 0.92

D. KD1 vs. KD3, IgG, accuracy = 0.84
Figure 3.
Figure 4.

A. KD1 vs. FC, training set and KD1 vs. FC, testing set.

B. Box plots showing intensity for various proteins across FC and KD1 conditions.
Figure 5.

A. KD3 vs. NC, training set vs. KD3 vs. NC, testing set

B. Box plots showing intensity distributions for various proteins in KD3 vs. NC.
Figure 6.